(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 December 2006 (07.12.2006)

PCT

(10) International Publication Number
WO 2006/130669 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2006/021088

(22) International Filing Date: 31 May 2006 (31.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/686,000	31 May 2005 (31.05.2005)	US
60/711,497	26 August 2005 (26.08.2005)	US
60/800,496	15 May 2006 (15.05.2006)	US

(71) Applicant *(for all designated States except US)*: **ENSEMBLE DISCOVERY CORPORATION** [US/US]; 99 Erie Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **STERN, Andrew, M.** [US/US]; 252 Commonwealth Avenue, Unit #5, Boston, MA 02116 (US). **HARBESON, Scott, L.** [US/US]; 203 Pemberton Street, #5, Cambridge, MA 02140 (US). **DRIGGERS, Edward, M.** [US/US]; 3 Huntington Road, Arlington, MA 02474 (US). **SPEAR, Kerry, L.** [US/US]; 601 Lexington Road, Concord, MA 01742 (US).

(74) Agents: **GREENHALGH, Duncan, A.** et al.; Goodwin Procter Llp, Exchange Place, Boston, MA 02109 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2006/130669 A2

(54) Title: ANCHOR-ASSISTED FRAGMENT SELECTION AND DIRECTED ASSEMBLY

(57) Abstract: The invention provides methods for compound and lead generation and discovery. In particular, the present invention provides a method for generating compounds and for selecting compounds that bind to a target. The present invention provides a way by which anchors (e.g., weak binders) and anchor-scaffold conjugates can be evolved into new generations of compounds having improved target binding and other desired pharmaceutical properties through control of both synthetic input and selection criteria.

ANCHOR-ASSISTED FRAGMENT SELECTION AND DIRECTED ASSEMBLY

RELATED APPLICATIONS

[0001] This application claims the benefit of and priority to U.S. Patent Applications Serial Nos. 60/686,000, filed on May 31, 2005; 60/711,497, filed on Aug. 26, 2005; and 60/800,496, filed on May 15, 2006, the entire disclosure of each of which is incorporated by reference herein for all purposes.

FIELD OF THE INVENTION

[0002] The present invention relates generally to DNA programmed chemistry and generation and discovery of compounds for target binding. More particularly, the present invention relates to methods for making and identifying organic molecules for binding to biological targets through anchor and/or fragment-based nucleic acid-templated chemistry.

BACKGROUND

[0003] Although chemistry and screening throughputs have increased significantly recently, drug lead discovery and development remain a high-risk, low-return process. An initial task in the generation of novel, biologically effective molecules is to identify and characterize binding ligands for a given biological target molecule. To date, this continues to be a daunting task in drug lead discovery. While many millions of compounds have been synthesized and screened, few have led to optimized compounds that eventually meet all the requirements of a drug.

[0004] More recently, fragment-based approaches for compound discovery have started to emerge. Small, diverse and information-rich fragments may provide more chemical space for optimization. Moreover, fragments of low complexity may be more likely to match a target binding site. As a result, certain compounds may still provide good starting points for optimization. Examples of such approaches include the "SAR by NMR" approach developed by Fesik *et al.* (U.S. Patent No. 5,698,401 by Fesik *et al.*; Shuker, *et al.*, 1996, Science, vol. 274, pp. 1531-1534), the "tethering" approach pioneered by Wells, *et al.* (U.S. Patent No. 6,335,155 by Wells, *et al.*; Erlanson, *et al.*, 2000, PNAS, vol. 97(17), pp. 9367-9372), a high-throughput x-ray crystallography method by Carr *et al.* (Carr, *et al.*, 2002, Drug Discovery Today, vol. 7, pp. 522-

527), and the use of surface plasmon resonance developed by Vetter *et al.* (Vetter, J., 2002, Cell. Biochem. Suppl., vol. 39, pp. 79-84).

[0005] In a manner analogous to the method of pharmacophore recombination (see, U.S. Patent No. 6,344,334 by Ellman *et al.*), these methods identify fragments that bind to biological targets of interest and then elaborate them into novel structures with greater affinity for the target. The structure-based methods then apply knowledge of the fragments bound to the binding site to the design of new ligands. Reactive functional groups on the fragments are utilized in pharmacophore recombinations to enable chemical assembly of the identified fragments in a combinatorial fashion to produce a library of new ligands that may have greater affinity for the target. The desired outcome of these methods is the identification of a drug lead compound that binds to a biological target of therapeutic interest.

[0006] These methods, however, suffer from several deficiencies. One is the requirement for large amounts of protein for use in the required structural studies (either X-ray or NMR). Relatively large amounts of target protein are also required for the biological screen required to test each fragment member individually for its ability to inhibit or bind the target. Because of the biological screening requirement, another issue is the requirement for fragments that are not only soluble but also well behaved under the assay conditions in the 10 μM to 1 mM (or higher) range. At these high concentrations, non-specific effects such as aggregation of the fragment molecules can yield erroneous or misleading results.

[0007] US Patent No. 6,335,155 describes a method for hit discovery that employs a covalent bond (a disulfide bond) to form a target/ligand conjugate in order to facilitate identification of organic ligands. This "tethering approach" is similarly used in US Patent No. 6,811,966 and US2002/0150947. These methods, however, suffer from several deficiencies. One is the requirement of the identification of a reactive group on the target molecule (or the introduction of a reactive group) that can be used to form a covalent bond with a ligand. Structural information of the target is therefore necessary. Another limitation is that the covalent bond between the target protein and the ligand limits screening to only a small area adjacent to the covalent bond, thereby leaving other areas of potential binding sites unexplored. Furthermore, the need for a disulfide bond limits the diversity of ligands that may be screened by these methods.

[0008] In another approach, self-assembling chemical libraries have been reported where such libraries are used for the identification of molecules for target binding. Organic molecules are linked to individual oligonucleotides that mediate the self-assembly of the library and provide a code associated with the organic molecules. See, e.g., U.S. Patent Application Publication No. 2004/0014090 A1 by Neri *et al.* and PCT International Publication No. WO 03/076943 A1.

[0009] While these and other approaches have provided additional tools for compound discovery, there is still a need for a more efficient and effective way of generating and selecting compounds for various pharmaceutical and other needs.

SUMMARY OF THE INVENTION

[0010] The present invention is based, in part, upon the discovery that nucleic acid-templated chemistry can be applied to compound and drug lead discovery in a way that greatly increase the efficiency of compound and drug lead generation and discovery. In particular, the present invention provides a unique way of generating drug-like compounds and selecting compounds for target binding. The present invention further provides a way by which compounds (e.g., compounds of low complexity) and compound fragments can be evolved from initial fragments into new generations of compounds having improved target binding and other desired pharmaceutical properties through control of both synthetic input and selection criteria. The present invention further provides a way by which anchors (e.g., weak binders) and anchor-scaffold (or -fragment/building blocks) conjugates can be evolved into new generations of compounds having improved target binding and other desired pharmaceutical properties through control of both synthetic input and selection criteria.

[0011] In the methods described herein, a nucleic acid molecule functions not only as a detection strand for identification of fragments that bind to a target but also templates the chemical assembly of those fragments (e.g., in a directed combinatorial approach) to achieve combinations of fragments into ligands of enhanced affinity. Fragment selection and directed assembly by nucleic acid-templated chemistry permits the identification of pharmacophores and their subsequent assembly into novel ligands with high affinity for the target. Unlike other methods that require each fragment molecule to be assayed individually, the methods of the present invention allow selection of fragment libraries, identification of multiple fragments simultaneously, and determination of the relative affinities of the fragments, which provides

structure-activity relationship (SAR) data that can be used in the design of the building blocks for use in the subsequent fragment assembly.

[0012] In one aspect, the invention provides a method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule. A target molecule is combined with a plurality of pre-selected test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to (i.e., does not hybridize to) the nucleotide sequence associated with other test molecules. A target binding element is harvested that binds to the target molecule binding site with a K_D of 10 mM or lower. The sequence of the oligonucleotide associated with the target binding element harvested is determined so as to identify the target binding element that binds with a K_D of 10 mM or lower. In one embodiment, the oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding element that binds with a K_D of 10 mM or lower. In this method, each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

[0013] In another aspect, the invention provides a method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule. The target binding elements so identified bind with a K_D of 10 mM or lower. A target molecule is combined with a plurality of pre-selected test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing (i.e., or does not hybridize) to the nucleotide sequences associated with other target binding elements. A target binding element is harvested that binds to the target molecule with a K_D of 10 mM or lower. The oligonucleotide associated with the target binding element harvested is amplified.

The sequence of the amplified oligonucleotide is determined so as to identify the target binding element having a K_D with the binding site of 10 mM or lower.

[0014] In yet another aspect, the invention provides an *in vitro* method for producing a molecule that binds to a pre-selected target molecule. The pre-selected target molecule includes a binding site that includes a first binding domain and a second binding domain. A template and a reagent are provided. The template includes a first target binding element attached to a first oligonucleotide that defines a first codon sequence. The first target binding element has a first K_D with the first binding domain of the binding site. The reagent includes a second target binding element attached to a second oligonucleotide that defines a first anti-codon sequence capable of hybridizing to the codon sequence. The second target binding element has a second K_D with the second binding domain. The template and the reagent are combined under conditions to permit the first codon sequence to hybridize to the first anti-codon sequence so as to bring the first and second target binding elements into reactive proximity. The first and second target binding elements are chemically coupled (e.g., in the absence of a ribosome) to produce a reaction product that binds to the preselected target molecule. In an embodiment, the reaction product has a K_D with the binding site less than (i) the first K_D of the first target binding element with the first binding domain, and (ii) the second K_D of the second target binding element with the second binding domain.

[0015] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of substantially all of the test molecules includes a target binding element associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other target binding elements.

[0016] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of at least some of the test molecules includes two or more target binding elements and is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0017] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of substantially all of the test molecules comprises two or more target binding elements and is associated with a corresponding oligonucleotide. The nucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0018] In yet another aspect, the invention provides a complex of a target molecule bound to a test molecule. The test molecule includes two or more target binding elements. The test molecule is associated with a corresponding oligonucleotide that has a nucleotide sequence that (i) identifies the test molecule and (ii) contains an amplification sequence. Each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

[0019] In yet another aspect, the invention provides a composition that includes a plurality of complexes. Each complex includes a target molecule bound to a test molecule. The test molecule includes two or more target binding elements. Each test molecule is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules. Each of substantially all of the target binding elements is linked to a functional group through which the target binding element is attached to the oligonucleotide.

[0020] In yet another aspect, the invention provides a composition that includes a plurality of complexes. Each complex includes a target molecule bound to a test molecule that includes two or more target binding elements. Each test molecule is associated with a corresponding oligonucleotide that has a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences of other test molecules.

[0021] In yet another aspect, the invention provides a method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule. The target binding elements so identified bind with a K_d of 10 mM or lower. A target molecule is combined with a plurality of pre-selected test molecules under conditions that permit

a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing (i.e., or does not hybridize) to the nucleotide sequences associated with other target binding elements. A target binding element is harvested that binds to the target molecule with a K_d of 10 mM or lower. The oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding element having a K_d with the binding site of 10 mM or lower.

[0022] In yet another aspect, the invention provides an *in vitro* method for producing a molecule that binds to a pre-selected target molecule. The pre-selected target molecule includes a binding site that includes a first binding domain and a second binding domain. A template and a reagent are provided. The template includes a first target binding element attached to a first oligonucleotide that defines a first codon sequence. The first target binding element has a first K_d with the first binding domain of the binding site. The reagent includes a second target binding element attached to a second oligonucleotide that defines a first anti-codon sequence capable of hybridizing to the codon sequence. The second target binding element has a second K_d with the second binding domain. The template and the reagent are combined under conditions to permit the first codon sequence to hybridize to the first anti-codon sequence so as to bring the first and second target binding elements into reactive proximity. The first and second target binding elements are chemically coupled (e.g., in the absence of a ribosome) to produce a reaction product that has a K_d with the binding site less than (i) the first K_d of the first target binding element with the first binding domain, and (ii) the second K_d of the second target binding element with the second binding domain.

[0023] In yet another aspect, the invention provides a method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule. A target molecule is combined with a plurality of test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to (i.e.,

does not hybridize to) the nucleotide sequence associated with other test molecules. A target binding element is harvested that binds to the target molecule binding site with a K_d of 10 mM or lower. The sequence of the oligonucleotide associated with the target binding element harvested is determined so as to identify the target binding element that binds with a K_d of 10 mM or lower. In one embodiment, the oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding element that binds with a K_d of 10 mM or lower. In this method, each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

[0024] In yet another aspect, the invention provides a method for identifying a target binding element capable of binding to a target molecule. A target molecule is combined with a plurality of test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to (i.e., does not hybridize to) the nucleotide sequence associated with other test molecules. A target binding element is harvested that binds to the target molecule binding site with a K_d of 10 mM or lower. The sequence of the oligonucleotide associated with the target binding element harvested is determined so as to identify the target binding element that binds with a K_d of 10 mM or lower. In one embodiment, the oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding element that binds with a K_d of 10 mM or lower. In this method, each of substantially all of the target binding elements has all of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

[0025] In yet another aspect, the invention provides a method for identifying a compound having a desired binding affinity to a target molecule. The method includes the following. A library is provided that includes a plurality of test compounds. Each of the test compounds includes (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence

informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. A reference compound is provided that includes the common binding moiety. The target molecule, the library of test compounds, and the reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule. The test compounds that exhibit greater binding affinity to the target molecule than the reference compound are harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby to identify the test compounds having a desired binding affinity to the target molecule.

[0026] In yet another aspect, the invention provides a method for identifying a compound having a desired binding affinity to a target molecule. The method includes the following. The target molecule, a plurality of test compounds, and a reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule. Each of the plurality of test compounds includes (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the associated test compound. The reference compound includes the common binding moiety. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The oligonucleotide sequences of the test compounds that bound to the target are determined.

[0027] In yet another aspect, the invention provides a method for detecting a second binding domain on a target molecule having a first binding domain. The method includes the following. A test compound is provided that includes (1) a first binding moiety having a binding affinity to the first binding domain of the target molecule, (2) a scaffold moiety connected to the first binding moiety through a bridging moiety, and (3) a defining oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the test compound. The first binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The effect of the test compound on the binding of a reference compound to the target molecule is determined. The reference compound comprises the first binding moiety. The data collected is analyzed to detect the presence of a second binding domain on the target molecule.

[0028] In yet another aspect, the invention provides a method for identifying a compound having a desired binding affinity to a target molecule. The method provides the following. A library is provided that includes a plurality of test compounds, wherein each of the test compound comprises (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The target molecule and the plurality of test compound are combined under conditions that permit binding of one or more of the plurality of test compounds to the target molecule if such test compounds with desired binding affinity are present. The test compounds bound to the target are harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby identifying the test compounds having a desired binding affinity to the target molecule.

[0029] In yet another aspect, the invention provides a method for selecting a compound having a desired binding affinity to a target molecule. The method includes the following. A library is provided that includes two subsets of test compounds. Each of the first subset of test compounds includes (1) a common binding moiety, (2) a first scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. Each of the second subset of test compounds includes (1) a second scaffold moiety, and (2) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The first scaffold and the second scaffold may be the same scaffold. A reference compound is provided that includes the common binding moiety. The target molecule, the library of test compounds, and the reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule. The test compounds that exhibit greater binding affinity to the target molecule than the reference compound are harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby to identify the test compounds having a desired binding affinity to the target molecule.

[0030] In yet another aspect, the invention provides a library of chemical compounds. The library includes a plurality of compounds. The compounds are prepared by one or more nucleic-acid-templated chemical reactions. Each of the compounds comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety. The first moiety has a dissociation constant of 10 mM or lower less to a binding domain of the target molecule.

[0031] In yet another aspect, the invention provides a compound. The compound comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety. The first moiety has a dissociation constant of 10 mM or lower less to a binding domain of the target molecule.

[0032] The foregoing aspects and embodiments of the invention may be more fully understood by reference to the following definitions, figures, detailed description and claims.

DEFINITIONS

[0033] The term, "anchor" as used herein, refers to a small molecule fragment, a small molecule or peptide having preselected binding affinity for a target, preferably (but not necessarily) with a molecular weight less than 250 daltons. An anchor may or may not contain further functionalization to facilitate subsequent DNA programmed chemistry.

[0034] The term, "amplification" or to "amplify", as used herein, relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art. (See, e.g., Dieffenbach, *et al.*, 1995, PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., pp. 1-5.).

[0035] It is contemplated, however, that amplification may not be necessary to conduct the methods of the present invention where the oligonucleotide sequences of interest (e.g., those that identify target binding elements and members of chemical libraries synthesized by DNA programmed chemistry) may be determined by methods that do not require amplification of the sequences (e.g., direct sequencing). Thus, where herein amplification is described as a step or a process, sequencing without prior amplification of an oligonucleotide is also contemplated.

[0036] The term, "associated with" as used herein describes the interaction between or among two or more groups, moieties, compounds, monomers, etc. When two or more entities are "associated with" one another as described herein, they are linked by a direct or indirect covalent or non-covalent interaction. Preferably, the association is covalent. The covalent association may be, for example, but without limitation, through an amide, ester, carbon-carbon, disulfide, carbamate, ether, thioether, urea, amine, or carbonate linkage. The covalent association may also include a linker moiety, for example, a photocleavable linker. Desirable non-covalent interactions include hydrogen bonding, van der Waals interactions, dipole-dipole interactions, pi stacking interactions, hydrophobic interactions, magnetic interactions, electrostatic interactions, etc.

[0037] The term, "bind" or "binding" as used herein in connection with the interaction between a target (e.g., a protein) and a potential binding compound indicates that the potential binding compound associates with the target to a statistically significant degree as compared to association with similar targets (e.g., proteins) generally (i.e., non-specific binding). Thus, a compound binds to a target when the compound has a statistically significant association with a target molecule. Preferably a binding compound interacts with a specified target with a dissociation constant (K_D or K_d) of 10 mM or less. A binding compound can bind with "extremely low affinity" (1 mM < K_D < 10 mM), "very low affinity" (100 μM < K_D < 1 mM), "low affinity" (10 μM < K_D < 100 μM), "moderate affinity" (1 μM < K_D < 10 μM), "moderately high affinity" (100 nM < K_D < 1 μM), or "high affinity" (K_D < 100 nM, e.g., K_D < 50 nM or 20 nM, or "very high affinity" (1 nM or sub-nanomolar < K_D < 10 nM)) depending on the dissociation constant.

[0038] The term, "binding site" as used herein, refers to an area on a target molecule that participate in molecular recognition by a binding compound. Binding sites embody particular shapes and often contain multiple binding domains (or "binding pockets") present within the binding site and collectively represent the binding site. By "binding domain" or "binding pocket" is meant a specific volume within a binding site. A binding domain can often be a particular shape, indentation or cavity in the binding site. Binding domains can contain particular chemical groups or structures that are important in the non-covalent binding of another molecule such as, for example, groups that contribute to ionic, hydrogen bonding, or van der

Waals interactions between the molecules. The binding site or domains may be known in advance, or discovered in the process of implementing the procedures described herein.

[0039] The term, "codon" and "anti-codon" as used herein, refer to complementary oligonucleotide sequences in a template strand and in a reagent (or transfer) strand, respectively, that permit the reagent strand to anneal to the template strand during DNA programmed chemistry. Codons on templates identify or encode the small molecules attached to the templates according to the reagents and/or target binding elements used and the chemical transformation performed. Anti-codons on reagent strands or a solid support interact through Watson-Crick base pairing with codons (i.e., specific sub-sequences within templates) in DNA programmed chemistry, thereby specifically delivering selected reagents (including, e.g., target binding elements) to the template in the DNA programmed chemistry process.

[0040] The term, "common binding moiety" as used herein, refers to an anchor moiety that is incorporated into an expanded molecule comprising the anchor moiety and a scaffold, fragment or building blocks.

[0041] The terms "complementary" as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base pairing. For example, the sequence "A-G-T" binds to the complementary sequence "T-C-A." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarily between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

[0042] The term, "detection strand" as used herein, refers to an oligonucleotide that includes a specific identification sequence and may include PCR primer binding sequences. The specific identification sequence identifies the fragment or molecule associated with the detection strand, and can be covalently attached via linker to a target binding elements. The specific identification sequence additionally is designed to ensure an absence of base-pairing with other detection strands.

[0043] The term, "K_D" or "apparent K_d" as used herein, refers to apparent dissociation constant as defined below.

$$K_d \text{ (or dissociation constant)} = \{[P]\cdot[L]\}/[P\cdot L]$$

where P is the target (e.g., protein) and L is a specific library member with the potential to bind to P.

$$K_D \text{ (or apparent dissociation constant)} = \{[P]_T \cdot (1 - \varepsilon \cdot N_{SB})\} / \varepsilon \cdot N_{SB}$$

where ε (or observed enrichment of L relative to all library members) = $\{[P \cdot L]/[L]_T\}/N_{SB}$;

N_{SB} is the non-specific background of total library bound in the absence of P expressed as a fraction of total library; $[P]_T$ represents total target concentration; $[L]_T$ represents the total specific ligand concentration. For $[P]_T >> [L]_T$, $[P] = [P]_T$, $[L]_T = [P \cdot L] + [L]$.

[0044] The term, "DNA programmed chemistry" (or "DPC") or "nucleic acid-templated chemistry" as used herein, refer to a method by which synthetic products are translatable into amplifiable information via oligonucleotide templates. Particularly, sequence specific control of chemical reactants to yield specific products is accomplished by (1) providing one or more templates, which have associated reactive units; (2) contacting one or more transfer units (reagents) having an anti-codon and reactive unit with one or more templates under conditions to allow for hybridization to the templates and (3) reaction of the reactive units to yield products (e.g., products being associated with an amplifiable template). The structures of the reactants and products need not be related to those of the nucleic acids of the template and transfer unit.

[0045] The term, "DPC-fragment" as used herein, refers to the molecular combination of a target binding element covalently linked to a nucleotide strand (e.g., via a linker) in such a way that the molecular combination can participate directly in a DPC process (and optionally also is functionalized for subsequent DPC processes). The nucleotide strand is a detection strand or a reagent strand that includes an anti-codon (selected to enable binding to a DPC template) and PCR primer binding sequences to enable amplification of the sequence.

[0046] The term, "hybridization" as used-herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

[0047] The term, "linker" as used herein, refers to any of a number of molecular entities (cleavable or non-cleavable) that can be used to covalently attach functionalized small molecules to their respective DPC reagent, template or detection strands.

[0048] The terms, "nucleic acid", "oligonucleotide" or "oligo" or "polynucleotide" as used herein refer to a polymer of nucleotides. The polymer may include, without limitation, natural

nucleosides (i.e., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine), nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, and 2-thiocytidine), chemically modified bases, biologically modified bases (e.g., methylated bases), intercalated bases, modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose), or modified phosphate groups (e.g., phosphorothioates and 5'-N-phosphoramidite linkages). Nucleic acids and oligonucleotides may also include other polymers of bases having a modified backbone, such as a locked nucleic acid (LNA), a peptide nucleic acid (PNA), a threose nucleic acid (TNA) and any other polymers capable of serving as a template for an amplification reaction using an amplification technique, for example, a polymerase chain reaction, a ligase chain reaction, or non-enzymatic template-directed replication.

[0049] The term "plurality" or "set" as used in a "plurality" or "set" of fragments or compounds is meant a collection of fragments or compounds. The fragments or compounds may or may not be structurally related. For example, the number of fragments or compounds may be anywhere from 10; 20; 50; 100; 1,000; 10,000; 100,000; 500,000; to 1,000,000 or more.

[0050] The term, "reagent strand" as used herein, refers to an oligonucleotide that include an anti-codon (and may include but does not require PCR primer sequences) that are associated with (e.g., covalently) a small molecule, which may be a target binding element, or any other molecular species that can participate in a DPC process.

[0051] The term, "reference compound" as used herein, refers to a compound that comprises the common binding moiety that retains the binding characteristics of the common binding moiety.

[0052] The term, "scaffold" as used herein, refers to a chemical compound having at least one site or chemical moiety suitable for functionalization. For example, a small molecule scaffold or molecular scaffold may have two, three, four, five or more sites or chemical moieties suitable for functionalization. These functionalization sites may be protected or masked as would be appreciated by a person of ordinary skill in the art. The sites may also be found on an underlying ring structure or backbone.

[0053] The term, "small molecule" as used herein, refers to an organic compound either synthesized in the laboratory or found in nature having a molecular weight less than 10,000 daltons, optionally less than 5,000 daltons, and optionally less than 1,500 daltons. Preferably, a small molecule has a molecular weight less than 1,000 daltons, optionally less than 500 daltons, and optionally less than 250 daltons.

[0054] The term, "target" as used herein, refers to any compound of interest, small molecule or polymeric, naturally occurring or non-naturally occurring, and biological molecules or otherwise. A target can be an enzyme, protein, peptide, carbohydrate, polysaccharide, glycoprotein, hormone, receptor, antigen, antibody, virus, substrate, metabolite, transition state analog, cofactor, inhibitor, drug, dye, nutrient, growth factor, cell, tissue etc., without limitation. For example, the binding region of a target molecule may include a catalytic site of an enzyme, a binding pocket on a receptor (e.g., a G-protein coupled receptor), a protein surface area involved in a protein-protein or protein-nucleic acid interaction (e.g., a hot-spot region), or a specific site on DNA (e.g., the major groove) or a site with no biological function. The natural function of the target could be stimulated (agonized), reduced (antagonized), unaffected, or completely changed by the binding depending on the precise binding mode and the particular binding site. A target can also be a surface of a material, e.g., the surface or coating of a polymeric material or a metallic material.

[0055] For example, a target and a small molecule having binding affinity toward the target may form a non-covalently interaction to associate the target with the binding molecule. Non-covalent binding includes the subsequent introduction of functional groups into the small molecule compound that causes covalent attachment to the target following the non-covalent molecular recognition and binding event.

[0056] Examples of targets include kinases, phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

[0057] The term, "target binding element" or "TBE" as used herein, refers to a molecule, e.g., a small molecule or peptide, a fragment, portion, framework or component thereof, that may participate in recognition and binding, for example, specific binding, to a particular target. The target binding element may bind to a binding domain of the binding domain of a target molecule.

[0058] For example, target binding elements may include small molecules or peptides with a molecular weight less than 250 daltons that may or may not have detectable affinity for a target (i.e. < or = 100 μM) using non-PCR based detection methods.

[0059] The target binding elements used typically represent fragments, structures, and/or frameworks found in known drugs or leads. Additionally, these target binding elements may be linked to functional groups that enable linkage to an oligonucleotide template. These target binding elements may be linked to additional functional groups to enable their subsequent use in DPC to build libraries of more elaborated molecules.

[0060] Examples of functionalization on target binding elements include glycine as a bi-functionalized methylene fragment for DPC; methylamine or acetic acid as analogous mono-functionalized fragments for DPC; para-aminobenzoic acid as a bi-functionalized benzene fragment for DPC; aniline or benzoic acid as analogous mono-functionalized fragments for DPC; glutamine as a bifunctionalized propionamide, etc.

[0061] Target binding elements may have various affinities toward a particular target. Target binding elements may bind to the target molecule with a K_D or K_d, e.g., less than 1 nM, 10 nM, 100 nM, 1 μM, 10 μM, 20 μM, 50 μM, 100 μM, 200 μM, 500 μM, 1 mM, 100 mM, 500 mM or 1 M or greater.

[0062] The term, "template" or "DPC template" as used herein, refers to a molecule including an oligonucleotide having at least one codon sequence suitable for DNA programmed chemistry (a template mediated chemical synthesis). The template optionally may include (i) a plurality of codon sequences, (ii) an amplification means, for example, a PCR primer binding site or a sequence complementary thereto, (iii) a reactive unit associated therewith, (iv) a combination of (i) and (ii), (v) a combination of (i) and (iii), (vi) a combination of (ii) and (iii), or a combination of (i), (ii) and (iii).

[0063] For example, a template may refer to an oligonucleotide that encodes the DNA programmed synthesis of a compound that contain elaborated target binding elements to be tested for target affinity. In this case, the template includes one or more codons that recruit reagents in the DPC process, as well as PCR primer regions, and may include specific endonuclease cleavage sites.

[0064] The term, "transfer unit" as used herein, refers to a molecule including an oligonucleotide having an anti-codon sequence associated with a reactive unit including, for example, a building block, monomer, monomer unit, molecular scaffold, or other reactant useful in DNA programmed chemistry (a template mediated chemical synthesis).

[0065] Throughout the description, where compositions are described as having, including, or comprising specific components, or where processes are described as having, including, or comprising specific process steps, it is contemplated that compositions of the present invention also consist essentially of, or consist of, the recited components, and that the processes of the present invention also consist essentially of, or consist of, the recited processing steps. Further, it should be understood that the order of steps or order for performing certain actions are immaterial so long as the invention remains operable. Moreover, unless specified to the contrary, two or more steps or actions may be conducted simultaneously.

BRIEF DESCRIPTION OF THE FIGURES

The invention may be further understood from the following figures in which:

[0066] FIG. 1 is a schematic representation of a target, binding site and binding domains in a binding site.

[0067] FIG. 2 is a schematic representation of target binding elements and corresponding DPC-fragments.

[0068] FIG. 3 is a schematic representation of an exemplary method for the discovery of target binding elements having binding affinities to a target.

[0069] FIG. 4 is a schematic representation of an exemplary method for assembly and selection of target binding elements for a target and modular iteration to refine target binding.

[0070] FIG. 5 is a schematic representation of an exemplary method for identification and selection of enriched and depleted target binding elements.

[0071] FIG. 6 is a schematic representation of one embodiment of an anchor-based approach for the identification of improved binding and novel binding sites and generation of compounds having binding affinities to such binding sites.

[0072] FIG. 7 is an exemplary set of oligonucleotide sequences useful for performing certain aspects of the present invention (presented on separate sheets).

[0073] FIG. 8 is a schematic representation of one embodiment of an anchor-based approach for the identification of improved binding and novel binding sites and generation of compounds having binding affinities to such binding sites.

[0074] FIG. 9 is a schematic representation of one embodiment of an anchor-based approach for the identification of drug hits and leads and novel binding sites.

[0075] FIG. 10 is a schematic representation of anchor conjugates.

[0076] FIG. 11 is a schematic representation of two exemplary architects of anchor conjugates.

[0077] FIG. 12 shows an example of an anchor conjugate involving macrocyclic fumaramides.

[0078] FIG. 13 is a schematic representation of an exemplary architect of a 3' DNA conjugate.

[0079] FIG. 14 is a schematic representation of an exemplary architect of a 5' DNA conjugate.

[0080] FIG. 15 lists exemplary target binding elements.

[0081] FIG. 16 is a schematic representation of an exemplary architect of DNA-fragment conjugated.

[0082] FIG. 17 is a schematic representation of a mix-and-split strategy for oligonucleotides and DPC fragments.

[0083] FIG. 18 shows an exemplary FOPP-labeled DPC fragment conjugate (and an anchor-fragment linked DNA conjugate).

[0084] FIG. 19 shows exemplary selections of anchor-based libraries against a biological target.

DETAILED DESCRIPTION OF THE INVENTION

[0085] The present invention provides a new approach to drug lead generation and selection where DNA programmed chemistry plays a critical role. Key attributes of DNA programmed chemistry that make such an approach possible and effective include: 1) the extreme sensitivity of PCR-linked binding assays to identify low affinity target binding elements, 2) the ability to test directly for binding in a manner that enables discovery of novel binding modes in novel fragment combinations, 3) the ability of DPC to rapidly assemble DPC-fragments into libraries of potentially high-affinity ligands, and 4) the modularity of the DPC system to allow rapid analysis and deconvolution of binding data from an entire library of compounds synthesized from DPC fragments.

[0086] The sensitivity of a PCR-based binding assay allows detection of low affinity interactions. Interactions in the range of 10 μM to 1 mM are difficult to detect by standard biochemical screening methods in which [Ligand] >> [Target]. Without wishing to be bound by theory, this may be due to the poor aqueous solubility of many small molecules and the tendency of some of these molecules to form aggregates in solution resulting in false positives. However, these affinity ranges may represent preferred starting points for hit to lead optimization. The PCR-based binding assays can detect the presence of as few as 1 DNA molecule and provide a basis for discovering target binding elements as DPC-fragments having affinities well within this affinity range. The use of target concentrations that exceed ligand concentrations is a central component of methods designed to detect low affinity binders – an inversion of the usual concentration requirements in an *in vitro* binding assay.

[0087] PCR-based binding assays may allow a method of detection that is independent of any specific target and independent of any target's biochemical activity. Selections of DPC fragments or compounds therefore employ a universal binding assay. The ability to screen exclusively for binding eliminates the requisite linkage to a functional biochemical assay; therefore, binding interactions can be detected that might otherwise fail to generate the functional biochemical readout. Selections can also be performed in the presence of soluble ligands for which the binding site of the ligand to the target is known. Under these conditions of increased stringency, knowledge regarding the binding of target binding elements to the target can be inferred. This approach uniquely enables the discovery of binding sites that lie outside the scope of interactions that provide a detectable biochemical output *in vitro*.

[0088] DPC enables the rapid assembly of DPC-fragments into potentially high-affinity compounds. DPC-fragments can be synthesized into compounds that may have high affinity to targets. In this novel fragment-based discovery approach, DPC-fragments identified can be assembled in a combinatorial fashion to yield libraries of more elaborated structures with an increased probability of providing moderate to high binding affinities (<< 10 μM). Other fragment-based approaches have no such facile method for converting identified fragments with low affinity into larger molecular weight compounds with high target affinity. In addition, the modular nature of DPC enables assembly of a variety of scaffolds and unstructured element display methods with equivalent synthetic ease, resulting in a variety of display options for the discovered target binding elements.

[0089] A fourth key advantage is the rapid analysis and deconvolution added by the modular nature of the data that comes from the target binding deconvolution process. The modularity of the DPC-fragment based system allows fast and efficient analysis and deconvolution of binding data from an entire library of compounds synthesized from DPC fragments. The sequence analysis of the identifying oligonucleotide sequence of a target binding fragment or molecule enables the rapid identification of its structure. When such data is acquired on a whole population of compounds (e.g., target binding fragments), the relative abundance of codons that are enriched (or depleted) among the binders can be compared to their relative abundance in the original library. The availability of such data that discretely links specific codons in the DPC-fragments with the affinity contribution of specific target binding elements in those compounds (and not just the overall compound affinity), on a library-wide scale, is a unique feature of a DPC-fragment approach. This data also facilitates iteration of the discovery cycle, with the possibility of re-using modular DPC reagents in subsequent cycles of syntheses, selections and analyses.

[0090] By employing the various components of DPC on the chosen fragments, from library synthesis to binding analysis, selection and evolution of the libraries, an efficient, unique, and superior method is hereby created for compound and drug lead discovery. The present invention permits the identification of pharmacophores and their subsequent assembly into novel ligands with high affinity for the target. For example, the fragment–based approach described herein allows identification of low molecular weight binders to target proteins that serve as viable

starting points for lead optimization. In addition, the present invention may be used in conjunction or combination with other methods of compound lead generation and discovery.

[0091] **FIG. 1** schematically illustrates a target **110**, one or more binding sites **210** and **220**, and binding domains in a binding site **310**, **320** and **330**.

[0092] There is generally no limitation as to the targets that may be investigated using the methods, compositions of matters and systems of the present invention. A target can be any compound of interest, small molecule or polymeric, and biological or otherwise. The target can be an enzyme, protein, peptide, carbohydrate, polysaccharide, glycoprotein, hormone, receptor, antigen, antibody, virus, substrate, metabolite, transition state analog, cofactor, inhibitor, drug, dye, nutrient, growth factor, cell, tissue etc., without limitation. Additional examples of biological targets include kinases, phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

[0093] **FIG. 2** schematically illustrates target binding elements **410**, **420**, **430** and **440** and corresponding DPC-fragments **510**, **520**, **530** and **540**. The DPC-fragments may contain a detection strand and/or a reagent strand.

[0094] Detection strands are designed to contain a primer binding sequence (for example, a 5' PCR primer binding sequence, a 3' PCR primer binding sequence, or both), and a specificity domain (e.g., a 4, 5, 6, 7, 8, or 10 base specificity domain). For sensitivity, the primer binding sites each include anywhere from 10 to 20 bases of sequence.

[0095] Criteria for designing the PCR primer binding sites include: 1) creating sufficient GC-content to allow annealing at an acceptable temperature, 2) minimizing palindromic sequences with respect to each other and within each primer binding site to avoid hairpin structures in the detection strand, and 3) minimization of reverse complementarity with any of the specificity domains.

[0096] Detection strands are introduced into a fragment-based discovery strategy by covalently attaching each of the strands to a pre-assigned TBE, through any of a variety of standard methods as described herein.

[0097] Detection strand sequences (including specificity domains) are designed according to the following exemplary scheme (for example using 6-mers, but can be anywhere from 4 to 20-

mers): (1) a list of all possible 6-mers is constrained to the set of sequences which have GC-content >1 and <5 (20%-80%, e,g., 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%), resulting in a set of 3200 sequences; (2) these sequences are included in exemplary detection strands; (3) an edge-node graph is generated with the resulting detection strands, where every sequence (node) is connected by an edge to every other sequence; (4) connections are eliminated where, for a given pair of nodes, there is a subsequence of length *S* or more in common between a node and the reverse complement of the other node. *S* may take on a variety of values, for example 3 or 4 or 5; (5) the resulting graph is analyzed for its "maximum cliques," which are the largest identifiable sets of nodes which are all completely inter-connected within the graph of 3200 nodes; (6) the resulting set of nodes in the maximum clique (for n = 6 base codons, and *S* = 4, a set of 510 such nodes can be found) represent detection strand sequences that are unlikely to form stable base-pairing structures between one another, and this expectation is confirmed using a standard oligo modeling program (e.g., OMP, produced by DNA Software Inc.).

[0098] A set of exemplary oligonucleotide sequences useful in performing the present invention are set forth in **FIG. 7**. Other examples of codon systems and detailed discussions can be found in Examples and in U.S. Patent Application Publication Nos. 2004/0180412 A1 by Liu *et al.* and 2003/0113738 A1, by Liu *et al.*

[0099] Reagent strand sequences are designed according to the strategy described above for designing specificity domains in order to minimize the degree of interaction between reagent strands and minimize base-pairing between unintended reagent strand and template codons and anti-codons. In addition to the specificity domain design elements, reagents may also contain fixed flanking sequences of 2-10 bases that act as registration domains that insure proper orientation of the specificity domains with the template. Reagent strand sequences typically do not contain PCR primer binding sequences, and the target binding elements are attached through cleavable linkers to enable DPC.

[0100] Various constraints are placed in the selection of fragments. The fragments are selected with a bias by compiling a set of known ligands/drugs for a particular type of targets and generating a set of fragments from these starting points based on the constraints. Libraries of know ligands and drugs can be compiled or synthesized based on publicly available information

and databases and are commercially available. Below in **Table 1** are examples of constraints that may be used in selecting fragments for target binding elements.

Table 1. Examples of Fragment Selection Constraints

Physical Property constraints	• Molecular Weight • ≤ cLogP • Number of Hydrogen Bond Donors (HBD) • Number of Hydrogen Bond Acceptors (HBD) • Polar Surface Area • Total Surface Area • Number of Rotatable Bonds

[0101] The constraints may be adjusted in both reactive functional groups and physical properties. For example, the molecular weight of the fragments may be constrained to be more than 90, 100, 110, 120, 150 daltons and less than 500, 450, 400, 350, 300, 250, 200, 150 daltons. The values of cLogP can be between -2 and 4, 5, 6, 7, 8, 9, or 10. The numbers of HBD and HBA can be 1, 2, 3, 4, 5, 6, 7 or be set to be more or less than any of these numbers. Polar surface area preferably is < 125 $Å^2$, more preferably < 100 $Å^2$, 80 $Å^2$, or 60 $Å^2$; total surface area preferably is <500 $Å^2$, more preferably < 400 $Å^2$, 300 $Å^2$, 200 $Å^2$ or 100 $Å^2$; the number of rotatable bonds preferably is <5, more preferably < 4 or 3. Other properties may be used as constraints as well such as the number of chiral centers, e.g., one or none; two of fewer; three or fewer chiral centers, etc.

[0102] Additional constraints that may be applied to fragment selection or synthesis are presence of certain functional groups that may be useful attaching fragments to oligonucleotide strands, as shown by non-limiting examples in **Table 2** below.

Table 2 Exemplary Functional Groups Useful to Attach Fragments to Oligonucleotide

Reactions	Functional Group Examples
5'-Amino Reacts with:	$-CO_2H$; –COCl; –NCO; –NCS; –OCOCl; –CHO; $-SO_2Cl$
5'-Amino Derivatized with Iodoacetyl Reacts with:	–SH
5'-Thiol Reacts with:	$COCH_2I$, Acrylamide, Maleimide, Epoxide
5'-Carboxy-NHS Ester Reacts with:	Amines
5'-Hydroxyl: Mitsunobu Reaction with:	Phenols, Imides
5'-Hydroxyl: Activation with Ms-Cl, Reacts with:	Amines, Thiols, Phenols, Imides, Stable carbanions

[0103] A library of DPC-fragments can include any number of members depending on the synthetic methods used to make the library and on the target to be investigated. For example, the fragment library may contain 100 or less, 500; 1,000; 5,000; 10,000 or more members.

[0104] Exemplary target binding elements have been identified for a number of targets. See, e.g., Erlanson, *et al.*, 2004, J. Med. Chem., vol. 47(14), pp. 3463-3482; Fattori, 2004, Drug Disc. Today, vol. 9(5), pp.229-239.

[0105] In one aspect, the invention provides a method for identifying a target binding element capable of *binding* to a binding domain disposed within a binding site of a target molecule. A target molecule is combined with a plurality of test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules. A target binding element is harvested that binds to the target molecule with a K_D with a binding site greater than 10 μM. The sequence of the oligonucleotide associated with the target binding element harvested is determined so as to identify the target binding element that binds with a K_D of 10 mM or lower. In one embodiment, the oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding

element that binds with a K_D of 10 mM or lower. In this method, each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

[0106] In another aspect, the invention provides a method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule. The target binding elements so identified have K_D values with the binding site greater than 10 μM. A target molecule is combined with a plurality of pre-selected test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule. Each test molecule includes a target binding element that is associated with an oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing (i.e., or does not hybridize) to the nucleotide sequences associated with other target binding elements. A target binding element is harvested that binds to the target molecule with a K_D greater than 10 μM. The oligonucleotide associated with the target binding element harvested is amplified. The sequence of the amplified oligonucleotide is determined so as to identify the target binding element having a K_D with the binding site greater than 10 μM.

[0107] In one embodiment, the method further includes the step of washing away unbound target binding elements after the combination of the plurality of pre-selected test molecules and harvest the target binding elements that bind to the target molecule with a pre-selected K_D, e.g., 1 μM, 10 μM, 20 μM, 50 μM or 100 μM. The method may further include washing away target binding elements that have a pre-selected K_D greater than, e.g., 50 μM, 100 μM, 200 μM, 500 μM, 1 mM, 100 mM, 500 mM or 1 M.

[0108] The target binding elements may have a mass ranging from 90 to 1,000 daltons. For example, the molecular weight of the target binding elements (e.g., fragments) may be constrained to be more than 90, 100, 110, 120, 150 daltons and less than 1,000, 500, 450, 400, 350, 300, 250, 200, or 150 daltons.

[0109] In one embodiment, the oligonucleotide is amplified by polymerase chain reaction wherein a primer anneals to the amplification sequence. A polymerase extends the primer annealed to the amplification sequence.

[0110] In yet another aspect, the invention provides an *in vitro* method for producing a molecule that binds to a pre-selected target molecule. The pre-selected target molecule includes a binding site that includes a first binding domain and a second binding domain. A template and a reagent are provided. The template includes a first target binding element attached to a first oligonucleotide that defines a first codon sequence. The first target binding element has a first K_D with the first binding domain of the binding site. The reagent includes a second target binding element attached to a second oligonucleotide that defines a first anti-codon sequence capable of hybridizing to the codon sequence. The second target binding element has a second K_D with the second binding domain. The template and the reagent are combined under conditions to permit the first codon sequence to hybridize to the first anti-codon sequence so as to bring the first and second target binding elements into reactive proximity. The first and second target binding elements are chemically coupled (e.g., in the absence of a ribosome) to produce a reaction product that has a K_D with the binding site less than (i) the first K_D of the first target binding element with the first binding domain, and (ii) the second K_D of the second target binding element with the second binding domain.

[0111] The method discussed here may include the step of selecting the reaction product. The method may further include the step of analyzing, e.g., by sequencing, the sequence of the first oligonucleotide associated with the reaction product. The sequence may also be determined by amplification. The sequence of the template is indicative of reaction product. The reaction product may include a first target element coupled to a plurality of second target elements.

[0112] In one embodiment, the first K_D of the first target binding element with the first binding domain is sufficient to permit the first target binding element to bind to the first binding domain in the absence of the second target binding element. In another embodiment, the first K_D of the first target binding element with the first binding domain is insufficient to permit the first target binding element to bind to the first binding domain in the absence of the second target binding element.

[0113] In another embodiment, the second K_D of the second target binding element with the second binding site is insufficient to permit the second target binding element to bind to the second binding domain in the absence of the first binding element.

[0114] In yet another embodiment, the first target binding element is known to bind to the first binding domain of the binding site. In one embodiment, the first target binding element is an anchor.

[0115] In one embodiment, the codon identifies the first target binding element associated with the first oligonucleotide. The anti-codon identifies the second target binding element associated with the second oligonucleotide. The template may include a plurality of different codons.

[0116] A plurality of different reagents may be combined with the template, and each reagent includes a different second target binding element attached to a corresponding, different oligonucleotide defining a corresponding anti-codon sequence. The anti-codon sequence is indicative of a particular second target binding element attached to the anti-codon.

[0117] **FIG. 3** schematically illustrates an exemplary method for the discovery of target binding elements that have binding affinities to a target. Target **110** having binding site **210** and domains **310**, **320** and **330** is combined with DPC-fragments **510**, **520**, **530** and **540** having target binding elements **410**, **420**, **430** and **440**, respectively. DPC-fragments **510** and **540** are harvested as they have the required binding characteristics (e.g., K_D). The corresponding oligonucleotide strands associated with **510** and **540** are amplified and deconvoluted to identify the DPC-fragments (revealing the identities of **510** and **540** which correspond to target binding elements **410** and **440**).

[0118] **FIG. 4** is a schematic representation of an exemplary method for assembly and selection of target binding elements for a target and modular iteration to refine target binding. Identified target binding elements **410**, **420**, **430**, **440**, etc., are assembled (e.g., by DPC) to create scaffolds **610**, **620**, **630**, **640**, **650**, etc. the assembly may be conducted under a pre-set criteria or randomly. The chemical assembly of the target binding elements can be accomplished using chemical methodologies that have been established as amenable to DPC. See, e.g., U.S. Patent Application Publication Nos. 2004/0180412 A1 and 2003/0113738 A1, Gartner *et al.*, 2004, Science, 305(10), pp. 1601-1605; Liu, *et al.*, 2002, Angew. Chem. Int. Ed., vol. 41(10), pp. 1796-2000). The TBE's can be linked directly to each other via covalent bonds or linker groups as shown for **610**, **620**, and **630** or they can be assembled using a scaffold. The scaffold can be flexible as in **640** or conformationally rigid as shown for **650**.

[0119] The new target binding elements (i.e., scaffolds) **610, 620, 630, 640, 650,** etc., are then subject to binding, oligonucleotide strand amplification and deconvolution so as to identify a subset of scaffolds that meet a certain binding characteristics (e.g., **630** and **650**). More rounds of re-combination and selection or screening can be carried out to apply higher or different stringencies to optimize for binding, selectivity and other properties. Structural analogs of the TBE's can also be incorporated into the additional rounds of the process to expand the SAR of the interactions at the target binding domain(s).

[0120] Selection and/or screening for desired activities (e.g., binding affinity, catalytic activity, or a particular effect in an activity assay) may be performed according to any applicable protocol. See, e.g., U.S. Patent Application Publication Nos. 2004/0180412 A1 by Liu *et al.* and 2003/0113738 A1, by Liu *et al.*

[0121] For example, affinity selections may be performed according to the principles used in library-based selection methods such as phage display, polysome display, and mRNA-fusion protein displayed peptides. Selection for catalytic activity may be performed by affinity selections on transition-state analog affinity columns (see, e.g., Baca *et al.*, 1997, Proc. Natl. Acad. Sci. USA 94(19): 10063-8) or by function-based selection schemes (see, Pedersen *et al.*, 1998, Proc. Natl. Acad. Sci. USA 95(18): 10523-8). Since minute quantities of DNA (about 10^{-20} mol) can be amplified by PCR (Kramer *et al.*, 1999, Current Protocols IN Molecular Biology (ed. Ausubel, F. M.) 15.1-15.3, Wiley), these selections can be conducted on a scale ten or more orders of magnitude less than that required for reaction analysis by current methods. The selection strategy does not require any detailed structural information about the target molecule or about the molecules in the libraries.

[0122] As schematically illustrated in **FIG. 5**, identification and selection of enriched and depleted target binding elements can be facilitated by the codons attached to the target binding elements.

[0123] In one embodiment, to allow deconvolution for DPC-fragments, DPC-fragments are designed to have only a single codon for identity, which renders the deconvolution process a relatively straight-forward analysis. Prior to a selection, the relative abundance of the various codons is determined by any of several methods, including real-time PCR (RT-PCR), microarray analysis, or single molecule sequencing. Following a selection, the same method is then applied, and the change in abundance of the DPC-fragment codons reveals enrichment or depletion. For

real-time PCR, a unique set of primers for each DPC-fragment are employed, each in a single PCR reaction is designed to amplify a particular codon. The unique primers will typically be comprised of a common PCR primer sequence plus a primer that recognizes the unique codon. Monitoring the crossing- threshold of each uniquely amplified sequence reveals the relative abundance of each component. For microarray analysis, a microarray must first be generated that contains the various sequences that are complementary to the full set of DPC-fragment codons. Using a two-color system where, for example Cy-3 is used to identify pre-selection, Cy-5 is used for post-selection. The relative Cy-3:Cy-5 ratio reveals the degree of enrichment. For single molecule sequencing, the relative abundance of each individual codon is determined directly from the abundance of a given sequence in the mixture pre- and post-selection.

[0124] In one embodiment, to allow deconvolution for products of DPC library synthesis, the same set of techniques can be used to reveal enrichment or depletion of DPC templates due to selection of DPC library components. However, the analysis must take into consideration that each unique sequence is composed of three codons, and that each individual codon will find itself in the context of multiple unique template sequences. One preferred method for deconvolution involves simply determining by RT-PCR the enrichment at the codon level. Then, evaluation of intramolecular chemical interactions reveals by codon-codon covariance in the raw enrichment data to identify the preferred total structures. It is important to note that a single distribution of codon frequencies does not uniquely determine the distribution of DPC library components. Similar data can also be acquired by microarray, or single molecule sequencing as described above. With these other techniques, codon-codon covariance again reveals intramolecular chemical interactions.

[0125] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of substantially all of the test molecules includes a target binding element associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other target binding elements.

[0126] In one embodiment, each of at least some of the target binding elements has a K_D with a target binding site greater than 10 μM. In another embodiment, each of substantially all of the target binding elements has a K_D with a target binding site greater than 10 μM. In another

embodiment, each of substantially all of the target binding elements has a molecular weight less than about 400 daltons.

[0127] In one embodiment, each of substantially all of the target binding elements is linked to a functional group through which the target binding element is attached to a corresponding oligonucleotide. Non-limiting examples of such functional groups include amines, carboxylic acids, acid chlorides, esters, ketenes, chloroformates, carbonates, aldehydes, acetals, thioacetals, ketones, ketals, thioketals, hydrazines, hydrazides, hydrazones, diazo compounds, esters, sulphonyl chlorides, alcohols, diols, phenols, azides, thiols, disulfides, isocyanates, isothiocyanates, alkyl and aryl halides, epoxides, aziridines, enamines, acrylamides, enones, maleimides, enolethers, imidates, oximes, nitrones, ylides, alkenes, dienes, and acetylenes.

[0128] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of at least some of the test molecules includes two or more target binding elements and is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0129] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of substantially all of the test molecules includes two or more target binding elements and is associated with an oligonucleotide. The nucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0130] A test molecule may include 2, 3, 4, 5, 6 or more target binding elements. Test molecules may have various affinities toward a particular target, e.g., with a K_D to a target molecule less than 1 nM, 10 nM, 100 nM, 1 μM, 10 μM, 20 μM, 50 μM, 100 μM, 200 μM, 500 μM, 1 mM, 100 mM, 500 mM or 1 M or greater.

[0131] In yet another aspect, the invention provides a complex of a target molecule bound to a test molecule. The test molecule includes two or more target binding elements. The test molecule is associated with an oligonucleotide that has a nucleotide sequence that (i) identifies the test molecule and (ii) contains an amplification sequence. Each of substantially all of the

target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons. As discussed herein, these and other constraints may be used to select target binding elements.

[0132] In yet another aspect, the invention provides a composition that includes a plurality of complexes. Each complex includes a target molecule bound to a test molecule. The test molecule includes two or more target binding elements, and each test molecule is associated with an oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules. Each of substantially all of the target binding elements is linked to a functional group through which the target binding element is attached to the oligonucleotide.

[0133] In yet another aspect, the invention provides a composition that includes a plurality of complexes. Each complex includes a target molecule bound to a test molecule that includes two or more target binding elements. Each test molecule is associated with an oligonucleotide that has a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences of other test molecules.

[0134] The anchor-based approach of the present invention employs a ligand (e.g., a pharmacophore) that is known or found to bind to a target and use it as an anchor to assist other potential pharmacophores bind to known or unknown target binding sites. Particularly in the anchor-based approach, by incorporating an anchor moiety into the library (e.g., of scaffolds or fragments), the apparent binding affinity of weak binders to a target can be increased, thus allowing them to be identified through selections.

[0135] **FIG. 6** is a schematic representation of one embodiment of anchor-assisted approach for identification of novel binding sites and generation of compounds having binding affinities to such binding sites. The fragments that are identified to have pre-selected binding properties can be optimized via a conventional medicinal chemistry approach, independent of amplifiable DNA conjugation, only if the method of observing binding is sufficiently sensitive to quantify weak interactions constituting an initial structure-activity relationship. Otherwise, the use of a high throughput ultra-sensitive DNA-dependent binding selection (e.g., Gartner *et al.*, 2004, Science,

vol. 305, pp1601-1605) is the method of choice. The latter method in conjunction with a DPC-based library approach, where one point of potential diversity on a particular scaffold is made invariant with the addition of a fragment, can be implemented. In this manner, the fragment serves as an "anchor" directing the library to the specific target of interest. Efficient selection of library members that bind more tightly to the target than the original anchor fragment alone provides a direct data-driven approach for lead optimization that is either independent of structural information for the target protein or can be complemented by it. Optimization of interactions distinct from the ones from the anchor fragment alone can lead to improved drug candidates analogous to the way the second generation ACE inhibitor, enalapril, was evolved from the first generation drug captopril.

[0136] In one embodiment of the anchor-based approach, as illustrated schematically in **FIG. 9**, an anchor **930** is chosen from known binders **910** or from a fragment library **920** via selection **930**. The anchor moiety is chemically incorporated (e.g., via DPC) at a point of diversity **950** in a library of compounds **960** (e.g., a diversity-oriented synthetic (DOS) DPC library) to generate an anchor-based subset of the original library (i.e., conjugates of the anchor moiety and the subset of the original library). A focused selection **970** is performed for the target of interest to which the selected anchor per se will bind to determine if positive selection is obtained for the members of the anchor-based subset (e.g., an anchor-based subset of the DOS DPC library). If positive selection is observed for the anchor-based subset resulting in a set of selected conjugates **980**, the selection can be tuned by adding varying concentrations of the corresponding non-conjugated anchor. The optimal concentration of competing anchor can be determined empirically. The selection is considered optimized (tuned) **985** when the positive selection for the members of the anchor-based subset is lowered to its limit of quantitative detection. This completes the selection of anchor.

[0137] The anchor-based subset, used as the training set, can now be expanded **950** into a larger chemically diverse anchor-based library **960**. The anchor moiety **940** (or an improved version) may now be incorporated into the larger library to generate an anchor-based library **960**.

[0138] Next, a selection **970** as tuned above, can now be performed to identify binders from the newly expanded anchor-based library **960** with affinities greater than the anchor per se. The stringency of the selection can be increased to enable the elucidation of SAR by decreasing the concentration of the target protein or by further increasing the concentration of the competing

anchor. The key point is that the higher affinity of certain library members will result from interactions at positions of diversity distinct from the anchor moiety.

[0139] The resulting SAR from the above selection of the expanded library can be used in the design of follow-up libraries. The above process may be iterated, and optimization of binding through this iterative process will enable the exploration of both novel chemical and biological space distinct from the original anchor moiety and its binding site on the target. In certain cases it may be appropriate to remove (lift) the original anchor moiety, allowing a closer study of new modes of binding and binding sites potentially addressing issues related to selectivity and other properties (e.g., mechanism-based and non-mechanism-based toxicity). See **FIG. 8**.

[0140] The anchor-based example above illustrates the use of an anchor to explore the target topology adjacent to the anchor binding site and to identify potentially new binding domains and small molecule pharmacophores for these domains. The anchor approach described herein does not require a covalent bond be formed between the anchor and the target of interest (i.e., without "tethering"). Thus, no structural knowledge about the target is necessary. This approach is complementary to the fragment approach disclosed herein that seeks to identify small molecules that bind with weak affinity to targets. One advantage of the invention is that it allows the anchor to direct pharmacophore exploration to a region of the target that has been shown to produce desired therapeutic effects through ligand binding. Binding of a ligand to a target in itself may be insufficient for a therapeutic effect; however, binding of a ligand to a target domain that elicits a desired therapeutic effect has a higher probability of success in drug discovery. This method enables a discovery platform that tightly and efficiently integrates chemistry and biology providing a direct means to identify totally novel structures with corresponding novel modes of binding action from known chemical and biological space.

[0141] The anchor-based approach may be implemented in various ways, as schematically illustrated in **FIG. 10**. In one approach **1010**, the oligonucleotide is linked directly to the anchor and not directly linked to the scaffold (or fragment or building blocks). As an example, Phg-Arylsulfonamide may be employed as an anchor to direct a macrocyclicfumaramide (MCF) library to the active site of carbonic anhydrase. In another approach **1020**, the oligonucleotide is directly linked to the scaffold (or fragment or building blocks) and not directly linked to the anchor. In yet another approach **1030**, the oligonucleotide is indirectly linked to both the anchor and the scaffold (or fragment or building blocks).

[0142] In another approach **1040**, the anchor may be an integral part of the scaffold and actually remains a part of the final optimized compound. In this approach, the anchor still functions to direct the fragment or scaffold to a binding domain of the target but also serves as an integral component of the resulting pharmacophore and continues in the iterative library process to yield the optimized moiety.

[0143] **FIG. 11** illustrates exemplary architects of anchor libraries. **FIG. 11(A)** and **(B)** show two alternative approaches in linking the anchor moiety and the diversity portion of the anchored compound. The total number of compounds may be controlled by the numbers of the anchor, attachment points, linkers, diversity building blocks, etc. Crystalline structures of the anchor and the target where available may be helpful in designing a library of compounds to address a particular target.

[0144] As an example of this approach, statine residues may be incorporated into a MCF library (see, e.g., U.S. Patent Application Publication Nos. 2004/0180412 A1 by Liu *et al.* and 2003/0113738 A1, by Liu *et al.*), **FIG. 12**. In this case, statine is a known moiety that can bind to the catalytic site of aspartyl proteases. By incorporating this residue into the MCF library at either R1, R2, or R3, the catalytic machinery is targeted with a known pharmacophore (anchor) and MCF members with appropriate topology for binding may be identified. In subsequent DPC library iterations, the anchor will remain and may also be optimized along with R2 and R3 (i.e. side chain diversity of statine). Although the statine residue may undergo structural changes in the optimization process, the overall topology of the MCF scaffold will remain intact and the modified anchor will be a part of the optimized molecules.

[0145] In one aspect, the invention provides a method for selecting a compound having a desired binding affinity to a target molecule. The method includes the following. A library is provided that includes a plurality of test compounds. Each of the test compounds includes (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. A reference compound is provided that includes the common binding moiety. The target molecule, the plurality of test compounds, and the reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to

compete for binding to the target molecule. The test compounds that exhibit greater binding affinity to the target molecule than the reference compound are harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby to identify the test compounds having a desired binding affinity to the target molecule.

[0146] In another aspect, the invention provides a method for identifying a compound having a desired binding affinity to a target molecule. The method includes the following. The target molecule, a plurality of test compounds, and a reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule. Each of the plurality of test compounds includes (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the associated test compound. The reference compound includes the common binding moiety. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The oligonucleotide sequences of the test compounds that bound to the target are determined.

[0147] In yet another aspect, the invention provides a library of chemical compounds. The library includes a plurality of compounds. The compounds are prepared by one or more nucleic-acid-templated chemical reactions. Each of the compounds comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety. The first moiety has a dissociation constant of 10 mM or lower to a binding domain of the target molecule.

[0148] In yet another aspect, the invention provides a method for detecting a second binding domain on a target molecule having a first binding domain. The method includes the following. A test compound is provided that includes (1) a first binding moiety having a binding affinity to the first binding domain of the target molecule, (2) a scaffold moiety connected to the first binding moiety through a bridging moiety, and (3) a defining oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the test compound. The first binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The effect of the test compound on the binding of a reference compound to the target molecule is determined. The reference compound comprises the first binding

moiety. The data collected is analyzed to detect the presence of a second binding domain on the target molecule.

[0149] In yet another aspect, the invention provides a method for identifying a compound having a desired binding affinity to a target molecule. The method provides the following. A library is provided that includes a plurality of test compounds, wherein each of the test compound comprises (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. The target molecule and the plurality of test compound are combined under conditions that permit binding of one or more of the plurality of test compounds to the target molecule if such test compounds with desired binding affinity are present. The test compounds bound to the target are harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby identifying the test compounds having a desired binding affinity to the target molecule.

[0150] In yet another aspect, the invention provides a method for selecting a compound having a desired binding affinity to a target molecule. The method includes the following. A library is provided that includes two subsets of test compounds. Each of the first subset of test compounds includes (1) a common binding moiety, (2) a first scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule. Each of the second subset of test compounds includes (1) a second scaffold moiety, and (2) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound. The first scaffold and the second scaffold may be the same scaffold. A reference compound is provided that includes the common binding moiety. The target molecule, the library of test compounds, and the reference compound are combined under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule. The test compounds that exhibit greater binding affinity to the target molecule than the reference compound are

harvested. The oligonucleotide sequences of the test compounds harvested are determined thereby to identify the test compounds having a desired binding affinity to the target molecule.

[0151] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of at least some of the test molecules includes two or more target binding elements and is associated with a corresponding oligonucleotide. The oligonucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0152] In yet another aspect, the invention provides a composition that includes a plurality of test molecules. Each of substantially all of the test molecules includes two or more target binding elements and is associated with an oligonucleotide. The nucleotide has a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

[0153] In yet another aspect, the invention provides a compound. The compound comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety. The first moiety has a dissociation constant of 10 mM or lower less to a binding domain of the target molecule.

[0154] The following examples contain important additional information, exemplification and guidance that can be adapted to the practice of this invention in its various embodiments and equivalents thereof. Practice of the invention will be more fully understood from these following examples, which are presented herein for illustrative purpose only, and should not be construed as limiting in anyway.

EXAMPLES

Example 1. Exemplary Anchors

[0155] Examples of anchors are shown in the following tables. **Tables 3 and 4** is a set of anchors that may be utilized for targets in the Carbonic Anhydrase class. Anchors for targets in the Kinase class, particularly BCR/Abl and VEGFR2, are shown in **Table 5**. Anchors for

phosphatase targets, particularly PTP1b, are shown in **Table 6.** These anchors can be prepared according to the general protocols described below.

Table 3. Carbonic Anhydrase Anchors: 5'-DNA Conjugates

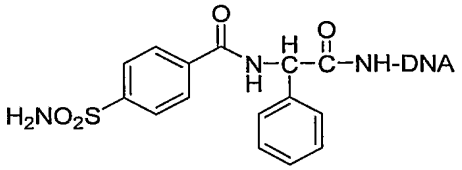
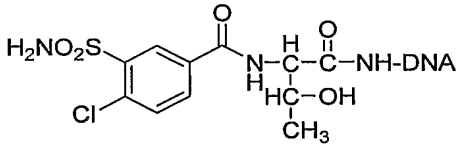
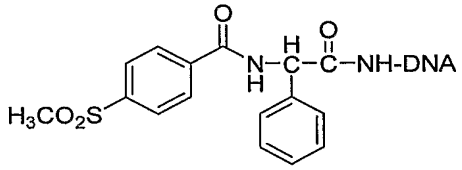
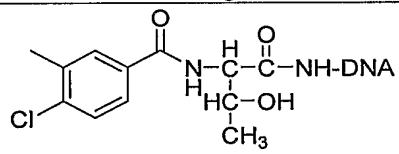
Anchor	Linker	Codons	Sequence	M^{-1} Digest (Calculated)
 "A-Phg"	5'-Amino-5	1-60-101	CAGACGTCA CGCCAAACT CACTACCAG CACTCTTCCG TCCACTACA AC (SEQ ID NO: 511)	709.1407 (709.1693)
 "B-Thr"	5'-Amino-5	29-60-101	CAGACGTCA CCAGAACCT CACTACCAG CACTCTTCCG TCCACTACA AC (SEQ ID NO: 512)	711.1437 (711.1253)
 "C-Phg"	5'-Amino-5	93-60-101	CAGACGTCA CAAGCCTCT CACTACCAG CACTCTTCCG TCCACTACA AC (SEQ ID NO: 513)	708.1271 (708.1741)
 "D-Thr"	5'-Amino 5	94-60-101	CAGACGTCA CTGTCCTCTC ACTACCAGC ACTCTTCCGT CCACTACAA C (SEQ ID NO: 514)	646.1626 (646.1681)

Table 4. Carbonic Anhydrase Anchors: 3'-DNA Conjugates

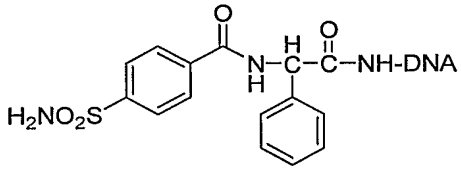
Anchor	Linker	Codons	Sequence	M^{-1} Digest (Calculated)
 "A-Phg"	3'-Amino C7	T40-6-45-85	CCACTACAA CACATCCCTC ACCCGTAAC ACTCCTTAGC CTCACCGCA ATCGAATTC CAC (SEQ ID NO: 515)	542.12 (542.14)

Table 5. Kinase Anchors: 3'-DNA Conjugates

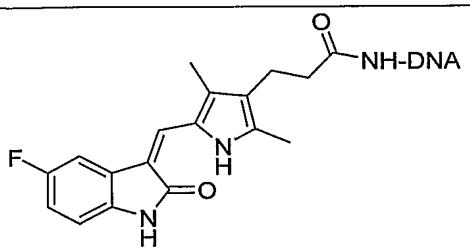
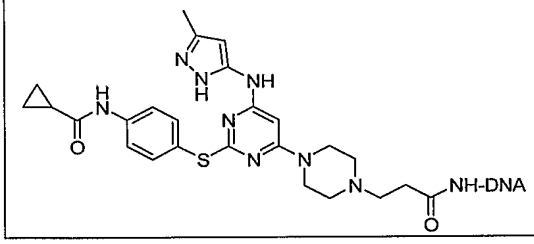
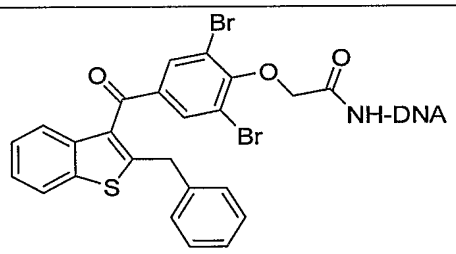
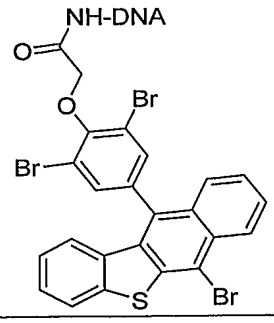
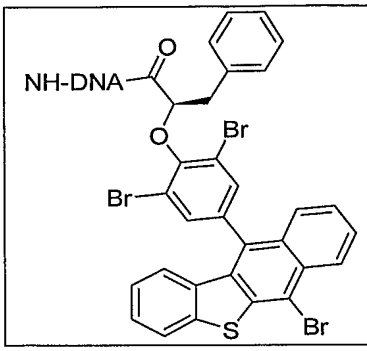
Anchor	Linker	Codons	Sequence	M^{-1} Digest (Calculated)
 "FOPP"	(mini-PEG2)-(miniPEG3)-3'-Amino C7	T40-3-42-65	CCACTACAACA CATCCCTCACC GTCAACACTCT ACAGCCTCACC GCAATCGAATT CCAC (SEQ ID NO: 516)	870.3602 (870.87)
	(mini-PEG2)-(miniPEG3)-3'-Amino C7	T40-25-56-109	CCACTACAACA CATCCCTCACC TCCTACACTCG CTTTCCTCACG ACCTTCGAATT CCAC (SEQ ID NO: 517)	1064.5043 (1064.47)

Table 6. Phosphatase Anchors: 3'-DNA Conjugates

Anchor	Linker	Codons	Sequence	MWt (Full Length)
	3'-Amino C7	T40-25-58-85	CCACTACAA CACATCCCTC ACCTCCTAC ACTCCCTAA GCTCACCGC AATCGAATT CCAC (SEQ ID NO: 518)	$(M-10)^{-10}$= 1843.4909
	3'-Amino C7	T40-25-58-86	CCACTACAA CACATCCCTC ACCTCCTAC ACTCCCTAA GCTCACCTG CATCGAATT CCAC (SEQ ID NO: 519)	$(M-10)^{-10}$= 1848.6775

	3'-Amino C7	T40-25-58-87	CCACTACAA CACATCCCTC ACCTCCTAC ACTCCCTAA GCTCACGCT CATCGAATT CCAC (SEQ ID NO: 520)	$(M-11)^{-11}$= 1688.7175

[0156] DNA oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8090 DNA synthesizer using standard phosphoramidite protocols and purified by reverse phase HPLC with a triethylammonium acetate/acetonitrile gradient. The 5'-amino modified oligo nucleotides were prepared by standard automated DNA synthesis using the 5'-amino-modifer 5 phosphoramidite from Glen Research. The 3'-amine modified oligonucleotides were prepared with the same protocol but with the 3'-amino-modifier C7 CPG from Glen Research. 3'-biotin oligonucleotides were prepared using Biotin TEG CPG from Glen Research.

[0157] To prepare the 5'-amine or 3'-amine modified anchored DNA strands, the oligonucleotides were prepared by standard automated DNA synthesis. The oligonucleotides were purified by RP-HPLC prior to conjugation to the various Anchor molecules. The general architectures of the 3'-amine or 5'-amine modified DNA strands are shown in **FIG. 13** and **FIG. FIG. 14**, respectively.

[0158] The anchor molecules as carboxylic acids were converted to the N-hydroxysuccinimide active esters, which were then conjugated to the 5'-amino or 3'-amino-modified oligos according to the following general protocols.

[0159] General protocol for preparing O-succinimidyl (OSu) ester: Free acid (0.5 mmole, 1 equiv.) and N-hydroxy succimide (0.6 mmole, 1.2 equiv.) were dissolved in 1.7 mL of anhydrous DMF under Ar, then N,N'-dicyclohexylcarbodiimide (DCC, 0.5 mmole, 1 equiv.) in 0.8 mL of anhydrous DMF were added (final concentration for free acid is 0.2 M). The reaction mixture was stirred at 40 °C for one to four hours. The extent of the OSu ester formation can be monitored by TLC. The presence of small amount of free acid can be neglected. After the reaction mixture was cooled in refrigerator (2 to 8 °C) for several hours, the precipitated

dicyclohexylurea (DCU) was then removed by filtration. The filtrate was treated with 15 mL of ether. Solid precipitated was washed three times with 10 mL of ether and dried under vacuum for several hours to afford the desired product (yield rang from 70 to 100 %).

[0160] General protocol for derivatizing DNA using OSu ester: To a 1.5 mL of centrifugation vial containing 50 nmole of DNA was added 104 μL of 0.1 M sodium phosphate buffer (NaPi), pH 8.6, 104 μL of OSu ester in NMP (96 mM or 72 mM) and 104 μL of NMP (final concentration for DNA: 0.16 mM). The vial was placed in a shaker and shaked at 37 °C for 1 hr to overnight. The extent of the DNA labeling can be monitored by analytical HPLC. The reaction mixture was desalted by gel filtration using Sephadex G-25 and then further purified by semi-preparative reversed-phase C18 column.

Example 2 Exemplary Fragments

[0161] A set of fragments (see **FIG. 15**) are chosen according to the constraints of **Table 7** below and modified as needed. The fragments are selected with a bias by compiling a set of known ligands/drugs for BCR/Abl and related kinases and generating a set of fragments from these starting points based on the constraints of **Table 7**. Libraries of know ligands and drugs can be compiled from publicly available information and databases and are commercially available.

Table 7. Examples of Fragment Selection Constraints

	Exemplary Constraints A	Exemplary Constraints B
Reactive functional groups:	• Primary and secondary amines • Primary anilines • Carboxylic acids • Bifunctional reagents containing an amine and a carboxylic acid moiety	• Primary and secondary amines • Primary anilines • Carboxylic acids • Bifunctional reagents containing an amine and a carboxylic acid moiety
Physical Property constraints	• 90 < Molecular Weight < 500 • $-2 \leq$ cLogP ≤ 4 • Hydrogen Bond Donors ≤ 4 • Hydrogen Bond Acceptors ≤ 8	• 90 < Molecular Weight < 300 • $-2 \leq$ cLogP ≤ 4 • Hydrogen Bond Donors ≤ 3 • Hydrogen Bond Acceptors ≤ 6

[0162] The constraints may be adjusted in both reactive functional groups and physical properties. For example, the molecular weight of the fragments may be constrained to be more

than 90, 100, 110, 120, 150 daltons and less than 500, 450, 400, 350, 300, 250, 200 or 150 daltons. The values of cLogP can be between -2 and 4, 5, 6, 7, 8, 9, or 10. The numbers of HBD and HBA can be 1, 2, 3, 4, 5, 6, 7 or be set to be more or less than any of these numbers. Other properties may be used as constraints as well such as the number of chiral centers, e.g., one or non; two of fewer; three or fewer, etc.; The number of NO_2 groups, e.g., 0, 1, 2, 3, 4, or more or less than any of these numbers. Additionally, the polar surface area, the total surface area, and the number of rotatable bonds may be used to define and select fragments.

Example 3. Fragment-based DPC Discovery

[0163] Each of the fragments is coupled to a specific DNA detection strand or reagent strand, and purified according to standard methods. There are many methods available to one skilled in the art for coupling strands to TBE's. Methods and references to these procedures can be readily obtained from many advanced text in organic chemistry, such as Carey, F.A. and Sundberg, R.J., Advanced Organic Chemistry Fourth Edition, Parts A & B, Kluwer Academic/Plenum Publishers, 2000; or March, Advanced Organic chemistry, John Wiley & Sons, New York, Fourth Edition, 1992. Non-limiting exemplary linkages include: amides (e.g., Carey *et al.* Part B, pp. 172-179); ureas (e.g., March, pp. 1299), carbamates (e.g., March, pp. 1280), sulfonamides (e.g., March, pp. 1296), aminoalkyl via reductive amination of amines with aldehydes or ketones (e.g., Carey *et al.*, pp. Part B. pp. 269-270), thioethers (e.g., Carey *et al.*, pp. 158; March, pp. 1297), ethers via Mitusunobu (e.g., Carey *et al.*, pp. 153-154), and carbon-carbon bonds via carbanions (e.g., Carey *et al.*, pp. 39-47) Purification of the DPC fragments can be accomplished by a number of methods available to those skilled in the art, such as but not limited to reverse phase HPLC, ion exchange chromatography and electrophoresis.

Preparation of Sample DPC Fragments

[0164] DNA oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8090 DNA synthesizer using standard phosphoramidite protocols and purified by reverse phase HPLC with a triethylammonium acetate/acetonitrile gradient. The 5'-amino modified oligo nucleotides were prepared by standard automated DNA synthesis using the 5'-amino-modifer 5 phosphoramidite from Glen Research. 5'-Thiol oligonucleotides were obtained with the 5'-Thiol-Modifier C6 from Glen Research. The 3'-amine modified oligonucleotides were prepared with the same protocol but with the 3'-amino-modifier C7 Controlled Pore Glass (CPG) from

Glen Research. 3'-biotin oligonucleotides were prepared using Biotin Triethyleneglycol (TEG) CPG from Glen Research.

[0165] To prepare the 3'-amine modified DPC fragments, the Fmoc-amine protected Target Binding Elements shown in **FIG. 15** were coupled to the 3'-amino-modifier C7 CPG using standard coupling protocols for peptide synthesis (Carey, F.A. and Sundberg, R.J., Advanced Organic Chemistry Fourth Edition, Part B, pp. 172-179). The oligonucleotides were then prepared by standard automated DNA synthesis. The architecture of the DPC Fragments is shown in **FIG. 16**. The 3-amino-modifier C7 is shown linking the Fragment to the 3' end of the DNA strand. From 3' to 5', the sequence consists of a PCR primer region, followed by the Position 3 codon that identifies the fragment. The position 2 and 1 codons follow and are available for templating DPC with complementary reagent strands. Position 0 represents a codon that uniquely identifies each sub-pool such that re-use of codons at positions 1-3 in different tag pools is enabled. The 5'-terminus is a PCR primer region.

[0166] The mix and split strategy was used in preparing the oligos as shown in **FIG. 17**. The 3'-amino-modifer CPG derivatized with the appropriate Fmoc-protected amino acids were extended with the appropriate 3'-PCR primer sequence followed by the fragment specific codon to provide 48 distinct CPG products. These were then grouped into 4 groups of 12 representing the common Tag sequences shown in **FIG. 15**. Each of the 4 groups of 12 products were then mixed to provide 4 mixtures that were then split into 12 equal portions to provide 48 portions of CPG for further DNA synthesis. This same mix and split procedure was followed for codon 2 and codon 1. After the addition of the nucleotide sequences for codon 1 and the mix step, the 4 resulting mixtures were then split in half to yield 8 groups of CPG. These were then extended with the appropriate Position 0 codons followed by the 5'-PCR primer sequence. This provided the 8 unique Tag pools with 48 fragments in which Tags A&B contained 12 fragments, Tags C&D contained 12 fragments, etc. The codon sequences used in the mix and split synthesis are shown in **Table 8**.

Table 8: Codon sequences used in a mix and split synthesis

5'-PCR Primer-Position 0 Sequence

atz1_c001_TA	CCACTACAACGCCAAACTC (SEQ ID NO: 521)	5'-PCR Primer-Pool A
atz1_c002_TB	CCACTACAACGAGCAACTC (SEQ ID NO: 522)	5'-PCR Primer-Pool B
atz1_c008_TC	CCACTACAACCAACCACTC (SEQ ID NO: 523)	5'-PCR Primer-Pool C
atz1_c011_TD	CCACTACAACTCAGCACTC (SEQ ID NO: 524)	5'-PCR Primer-Pool D
atz1_c019_TE	CCACTACAACCTAGGACTC (SEQ ID NO: 525)	5'-PCR Primer-Pool E
atz1_c032_TF	CCACTACAACATCCACCTC (SEQ ID NO: 526)	5'-PCR Primer-Pool F
atz1_c039_TG	CCACTACAACTCTACCCTC (SEQ ID NO: 527)	5'-PCR Primer-Pool G
atz1_c080_TH	CCACTACAACTCTCTGCTC (SEQ ID NO: 528)	5'-PCR Primer-Pool H

Position 1 Sequence

atz1_c003_1A	ACCGTCAACAC (SEQ ID NO: 529)
atz1_c005_1B	ACCACGAACAC (SEQ ID NO: 530)
atz1_c006_1C	ACCCGTAACAC (SEQ ID NO: 531)
atz1_c017_1D	ACAACCGACAC (SEQ ID NO: 532)
atz1_c024_1E	ACGCACTACAC (SEQ ID NO: 533)
atz1_c025_1F	ACCTCCTACAC (SEQ ID NO: 534)
atz1_c027_1G	ACCCTGTACAC (SEQ ID NO: 535)
atz1_c037_1H	ACGAAACCCAC (SEQ ID NO: 536)
atz1_c038_1I	ACATGACCCAC (SEQ ID NO: 537)
atz1_c041_1J	ACTTCTCCCAC (SEQ ID NO: 538)
atz1_c012_1K	ACATCGCACAC (SEQ ID NO: 539)
atz1_c034_1L	ACACTGACCAC (SEQ ID NO: 540)

Position 2 Sequence

atz1_c042_2A	TCCATTCCCTC (SEQ ID NO: 541)
atz1_c044_2B	TCTACAGCCTC (SEQ ID NO: 542)
atz1_c045_2C	TCCTTAGCCTC (SEQ ID NO: 543)
atz1_c051_2D	TCTAGCTCCTC (SEQ ID NO: 544)
atz1_c052_2E	TCAGTCTCCTC (SEQ ID NO: 545)
atz1_c054_2F	TCAACGTCCTC (SEQ ID NO: 546)
atz1_c055_2G	TCCTGTTCCTC (SEQ ID NO: 547)
atz1_c056_2H	TCGCTTTCCTC (SEQ ID NO: 548)
atz1_c058_2I	TCCCTAAGCTC (SEQ ID NO: 549)
atz1_c060_2J	TCTACCAGCTC (SEQ ID NO: 550)
atz1_c064_2K	TCCTCTAGCTC (SEQ ID NO: 551)
atz1_c031_2L	TCTCACACCTC (SEQ ID NO: 552)

Position 3 Sequence-3'-Primer Sequence

atz1_c065_3A	ACCTAACGCGAATTCCAC (SEQ ID NO: 553)
atz1_c078_3B	ACCACATGCGAATTCCAC (SEQ ID NO: 554)
atz1_c085_3C	ACCGCAATCGAATTCCAC (SEQ ID NO: 555)
atz1_c086_3D	ACCTGCATCGAATTCCAC (SEQ ID NO: 556)
atz1_c087_3E	ACGCTCATCGAATTCCAC (SEQ ID NO: 557)
atz1_c088_3F	ACCCAGATCGAATTCCAC (SEQ ID NO: 558)
atz1_c101_3G	ACTTCCGTCGAATTCCAC (SEQ ID NO: 559)
atz1_c102_3H	ACCATCGTCGAATTCCAC (SEQ ID NO: 560)
atz1_c108_3I	ACCGACTTCGAATTCCAC (SEQ ID NO: 561)
atz1_c109_3J	ACGACCTTCGAATTCCAC (SEQ ID NO: 562)
atz1_c112_3K	ACCCCTTTCGAATTCCAC (SEQ ID NO: 563)
atz1_c049_3L	ACCCAATCCGAATTCCAC (SEQ ID NO: 564)

[0167] The identity of the DPC fragments was confirmed by LC/MS analysis. Prior to the digestion protocol, any basic primary or secondary amines was acetylated to facilitate negative ionization. A solution of the DPC fragments in 0.3M TEAA (triethylammonium acetate) buffer, pH 7.2, was prepared (approx. 100 pmol in 200 µL), which was then treated with acetic anhydride (2 µL) for 30 min at room temperature. These samples were then evaporated to dryness. Oligo analytes were dissolved in 10 µL 10% methanol, and 1 µL internal standard solution (#1 below), 1 µL 10x buffer (#3 below), and 1 µL prepared Nuclease S1 aqueous Solution (#2 below) was added. Mixing was performed in a 600-ul plastic vial and the mixture was incubated at 37 °C in an air incubator for 2 hours.

[0168] The digestion control internal standard solution (#1) was comprised of 0.5 pmol/µl 1µL A-phg-E stock solution (product m/z 709, 8.7 µM) and 1 µL (product m/z 896, 10 uM) stock solution mixed with 18 µL H_2O; Store this solution in -20°C. The 40 unit/µL enzyme solution (#2) was comprised of 1 µL commercial Nuclease S1 (Roche Diagnostics GMBH, 400unit/ul) mixed with 9 µL H_2O. This solution is made right before using. The 10x digestion buffer (#3) was comprised of 330mM sodium acetate, 500 mM naCl, 0.33mM $ZnSO_4$, pH 4.5.

[0169] The digested samples were analyzed on an LC-MS system that consisted of a UPLC and Q-TOF premier mass spectrometer (Waters Corporation, Milford, MA). An Acquity column 100mm x 1mm i.d. was installed and the samples were eluted using a gradient elution at 50ul/min from 95% mobile phase A to 50% in 45 min. (HPLC mobile phase A: 1% hexafluoropropanol, 0.1% triethylamine in H2O; mobile phase B: Methanol). Negative ions were analyzed with mass spectrometer.

[0170] The results of the LC/MS analysis of the DPC fragment examples are shown in **Tables 9-12.**

Table 9. Pool A&B Templates LC/MS Analysis

Compound	Formula	Expected mass	Min M/Z	Max M/Z
Phosphate-3AMC7-ArgMe2-Ac	C17H36N5O7P	452.2274	451.6394	452.6462
Phosphate-3AMC7-D-HoCit-Ac	C16H33N4O8P	439.1958	437.9133	441.0056
Phosphate-3AMC7-LysFor-Ac	C16H32N3O8P	424.1849	422.4642	426.8094
Phosphate-3AMC7-Valeram-Ac	C16H30N30O8P	422.1692	421.1893	424.558
Phosphate-3AMC7-Cxhca-Ac	C17H33N2O7P	407.1947	405.8442	409.9454
Phosphate-3AMC7-Acyptene-Ac	C15H27N2O7P	377.1478	375.2831	379.737
Phosphate-3AMC7-Gaba-Ac	C13H27N2O7P	353.1478	352.2337	354.8958
Phosphate-3AMC7-AMeProp-Ac	C13H27N2O7P	353.1478	352.1867	354.9552
Phosphate-3AMC7-Gln-Ac	C14H28N3O8P	396.1536	395.8165	397.6751
Phosphate-3AMC7-bGln-Ac	C14H28N3O8P	396.1536	395.8165	397.6751
Phosphate-3AMC7-Ser-Ac	C12H25N2O8P	356.13	Not Found	
Phosphate-3AMC7-D-Ser-Ac	C12H25N2O8P	356.13	Not Found	

Table 10. Pool C&D Templates LC/MS Analysis

Compound	Formula	Expected mass	Min M/Z	Max M/Z
Phosphate-3AMC7-LysAc-Ac	C17H34N3O8P	438.2005	437.8456	440.3896
Phosphate-3AMC7-Lys(Nic)-Ac	C21H35N4O8P	501.2114	500.9587	503.4807
Phosphate-3AMC7-Met(O2)-Ac	C14H29N2O9PS	431.1253	430.8351	433.2946
Phosphate-3AMC7-A4PyrBA-Ac	C18H30N3O7P	430.1743	429.8668	432.487
Phosphate-3AMC7-A3PyrBA-Ac	C18H30N3O7P	430.1743	429.8668	432.487
Phosphate-3AMC7-SA4PBA-Ac	C18H30N3O7P	430.1743	429.8668	432.487
Phosphate-3AMC7-THPO2Gly-Ac	C16H31N2O9PS	457.141	456.8624	459.402
Phosphate-3AMC7-ACHXA-Ac	C16H31N2O7P	393.1791	392.9446	395.4278
Phosphate-3AMC7-D3Pal-Ac	C17H28N3O7P	416.1587	415.934	418.2854
Phosphate-3AMC7-HoWSer(Me)-Ac	C14H29N2O8P	383.1583	383.0061	385.4017
Phosphate-3AMC7-MeHis-Ac	C16H29N4O7P	419.1696	418.9733	420.2447
Phosphate-3AMC7-Gly - Ac	C11H23N2O7P	325.1165	324.4351	327.3268

Table 11. Pool E & F Templates LC/MS Analysis

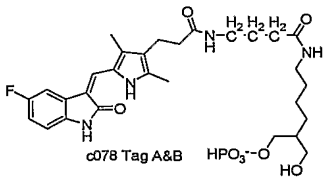
Compound	Formula	Expected mass	Min M/Z	Max M/Z
Phosphate-3AMC7-Val-Ac	C14H29N2O7P	367.1634	367.0273	369.3517
Phosphate-3AMC7-AFurBA-Ac	C17H29N2O8P	419.1583	418.9984	421.3014
Phosphate-3AMC7-Ala2Fur-Ac	C16H27N2O8P	405.1427	404.9481	407.3366
Phosphate-3AMC7-Ala4Thz-Ac	C15H26N3O7PS	422.1151	421.9503	424.2383
Phosphate-3AMC7-AMChxA -Ac	C18H35N2O7P	421.2104	420.9165	423.3477
Phosphate-3AMC7-AThiBA-Ac	C17H29N2O7PS	435.1355	434.6233	437.4323
Phosphate-3AMC7-AZPC -Ac	C21H32N3O8P	484.1849	483.6052	486.5442
Phosphate-3AMC7-CNHoPhe-Ac	C20H30N3O7P	454.1743	453.6125	456.4902
Phosphate-3AMC7-CypAla-Ac	C15H29N2O7P	379.1634	378.8961	381.3416
Phosphate-3AMC7-Dala4Thz-Ac	C15H26N3O7PWS	422.1151	421.8904	424.311
Phosphate-3AMC7-DiMeoPhe-Ac	C20H33N2O9P	475.1845	474.7141	477.539
Phosphate-3AMC7-L3Pal -Ac	C17H28N3O7P	416.1587	415.8744	418.4505

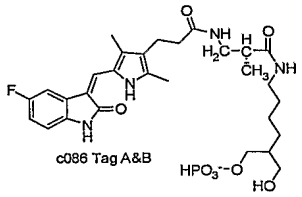
Table 12. Pool G & H Templates LC/MS Analysis

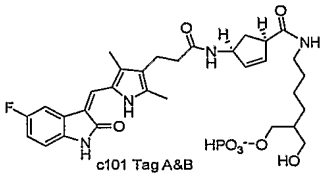
Compound	Formula	Expected mass	Min M/Z	Max M/Z
Phosphate-3AMC7-Cha-Ac	C18H35N2O7P	421.2104	421.0835	423.4081
Phosphate-3AMC7-ABztB-Ac	C21H31N2O7PS	485.1511	484.1497	488.0578
Phosphate-3AMC7-Thi-Ac	C16H27N2O7PS	421.1198	420.986	423.2295
Phosphate-3AMC7-DBip-Ac	C24H33N2O7P	491.1947	490.4798	495.0777
Phosphate-3AMC7-F2HoPhe -Ac	C19H29F2N2O7P	465.1602	464.7077	467.5954
Phosphate-3AMC7-Freidam-Ac	C19H36N3O8P	464.2162	463.9187	464.6658
Phosphate-3AMC7-His(Bn) -Ac	C22H33N4O7P	495.2009	493.9567	497.9117
Phosphate-3AMC7-Indanygly-Ac	C20H31N2O7P	441.1791	440.5078	444.1657
Phosphate-3AMC7-Styrylala-Ac	C20H31N2O7P	441.1791	440.4843	444.1423
Phosphate-3AMC7-LBip-Ac	C24H33N2O7P	491.1947	490.6873	495.0122
Phosphate-3AMC7-Leu-Ac	C15H31N2O7P	381.1792	380.5196	383.8154
Phosphate-3AMC7-Phg -Ac	C17H27N2O7P	401.1478	399.6447	404.5285

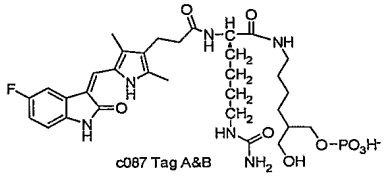
[0171] Example of DPC Assembly of Fragments. Pools A, E & H were shown to have very weak affinity for the target kinases, Abl & KDR. FOPP Target Binding Element was shown to have weak affinity for Abl and good affinity for KDR. DPC was used to assemble libraries that combined the FOPP Target Binding Element with the weak affinity DPC fragments exemplified above. Structures of compounds in Pools A, E & H are shown in **Tables 13-15.**

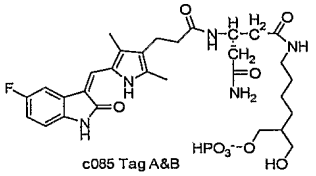
Table 13 Structures of Compounds in Pool A

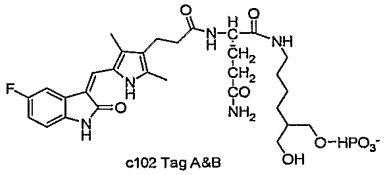

c078 Tag A&B

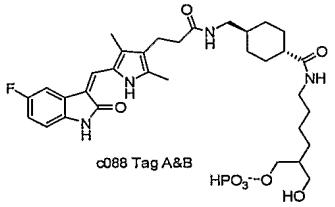

c086 Tag A&B

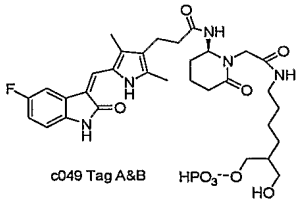

c101 Tag A&B

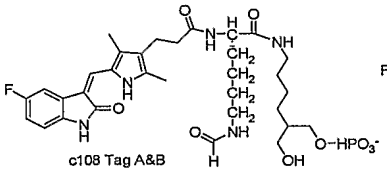

c087 Tag A&B

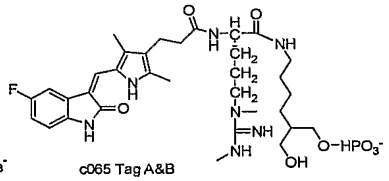

c085 Tag A&B

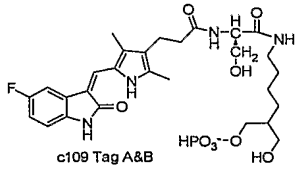

c102 Tag A&B


c088 Tag A&B


c049 Tag A&B


c108 Tag A&B


c065 Tag A&B


c109 Tag A&B

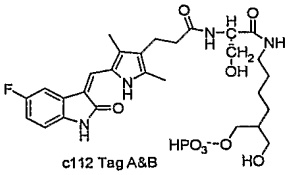
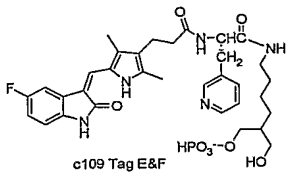
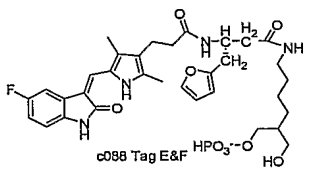
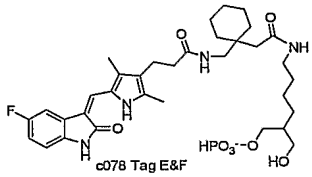
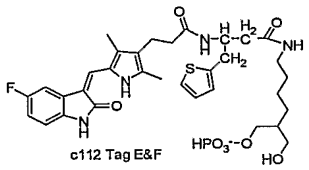

c112 Tag A&B

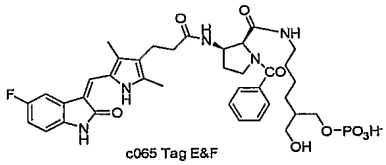
Table 14 Structures of Compounds in Pool E

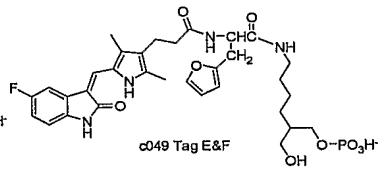

c109 Tag E&F

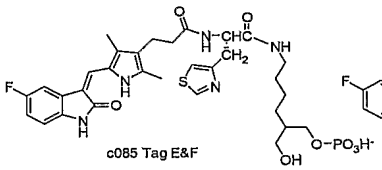

c088 Tag E&F

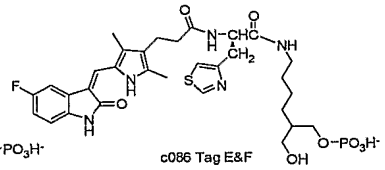

c078 Tag E&F

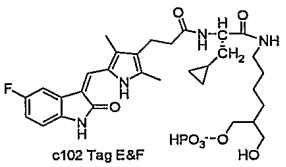

c112 Tag E&F

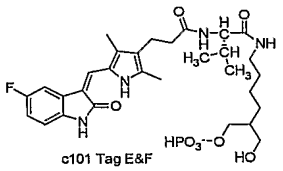

c065 Tag E&F

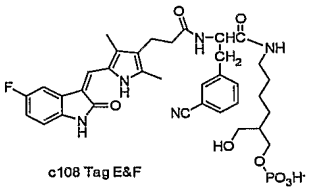

c049 Tag E&F


c085 Tag E&F


c086 Tag E&F


c102 Tag E&F


c101 Tag E&F


c108 Tag E&F

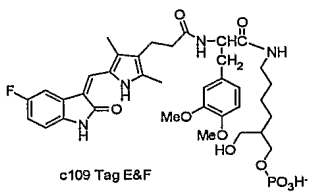
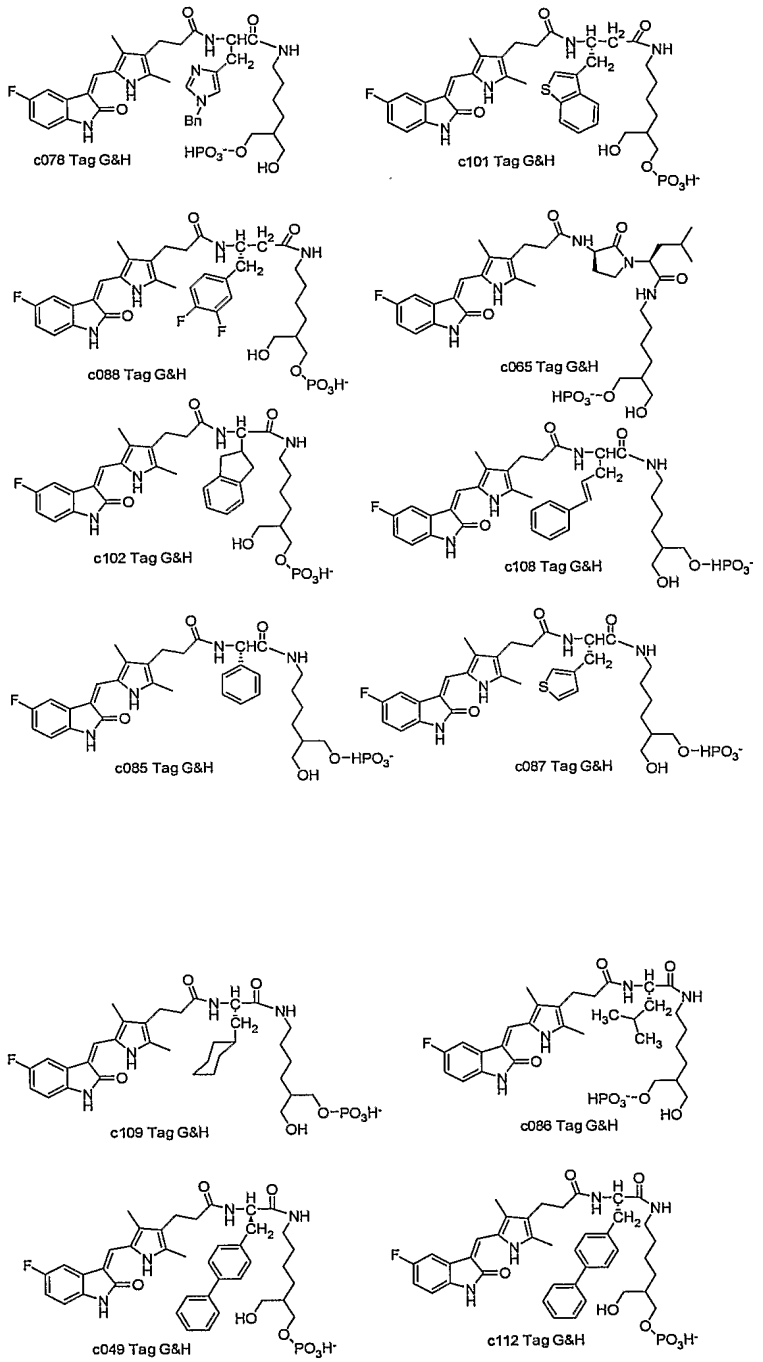

c109 Tag E&F

Table 15 Structures of Compounds in Pool H



[0172] Preparation of the DNA-FOPP Target Binding Element strand for DPC assembly of fragments. A general protocol for preparing DNA-OSu-R reagent can be found, for example, in

"Ordered Multistep Synthesis in a Single Solution Directed by DNA Templates" Snyder, T. M. and Liu, D. R. *Angew. Chem. Int. Ed.* **44**, 7379-7382 (2005). Briefly, a 10-mer DNA was prepared: 5'-trityl-S-GTG GAA TTC G-3'-biotin. The deprotection of the trityl group proceeded as follows. First, 40 μL of DNA (50-100 μM) was mixed with 2 μL 2.0 M TEAA (pH=7.0) and 6 μL $AgNO_3$ (1 M in H_2O) was added. The reaction was kept at RT on a vortexer for 30 min before 8.4 μL DTT (1 M solution in H_2O) was added and vortexed for 5 min to precipitate excess DTT. The yellowish suspension was loaded to a NAP 5 column and the 1 mL collected DNA solution was reacted with N-hydroxymaleimide in the next step. Next, 10 mg *N*-hydroxymaleimide was added with 125 μL H_2O and 125 μL MOPS (1M, pH=7.5). The solution turned brown immediately upon the addition of MOPS, and it was quickly mixed with the DNA solution obtained from last step. The reaction mixture was kept under RT for 30 min, then was placed in a speedvac to reduce the volume under 1 mL before being desalted on a NAP10 column. The product was purified by HPLC and then reacted with FOPP Target Binding element carboxylic acid. Then, 2.04 μmol FOPP-COOH was dissolved in 50 μL DMF, 0.5 mg EDC was dissolved in 50 μL DMF; and then 20 μL EDC solution, 25 μL FOPP-COOH solution and 5 μL DMF were combined. The reaction mixture was kept under RT for 20 min before being added to the DNA-solution (16 μL MES (0.5M, pH=6.5), 24 μL H_2O, and 40 μL DNA prepared in step 3.). The reaction was maintained at RT for 5-10 min, then desalted by a NAP5 column and purified by HPLC. After collecting the product fraction from the HPLC, 1:5 (v:v) 6% TFA was added directly into the fraction before putting it on the lyophilizer. The final product dried from TFA-containing lyophilization was yellow and in a semi-dry form, and was stored at -80°C in this form. Prior to use, the product was brought to 10-20 μM with H_2O, the concentration was measured, and then it was immediately used in DPC reaction. The structure of the FOPP-labeled DPC Fragment was confirmed by LC/MS (expected mass (6-): 710.6466; Observed mass (6-): 710.6719), as shown in **FIG. 18**.

[0173] DPC-Based Fragment Assembly. Assembly was performed under the following conditions: 1 M NaCl, 0.2 M MES (pH=6.5), 1 μM template, 2 μM DNA-FOPP reagent, at room temperature for 1 hour. The reaction was quenched with 1:20 (volume : volume) Tris-HCl buffer (1 M, pH=7.2), then subjected to a streptavidin tip purification to remove biotinylated reagents. The solution then was collected and dried in a speedvac until the volume was less than 0.5 mL to be desalted on a NAP5 column. The 1 mL solution collected from NAP5 column was

lyophilized and analyzed by LC-MS. The expected and observed molecular ions are shown in **Tables 16-18.**

Table 16. DPC Assembled Fragments LC/MS Analysis (Pool A)

Compound	Formula	Codon	Expected mass	Min. M/Z	Max. M/Z
Phosphate-3AMC7-Gaba_FOPP	C29H40FN4O8P	c078	621.249	620.3623	621.7786
Phosphate-3AMC7-AMeProp_FOPP	C29H40FN4O8P	c086	621.249	620.2646	621.7053
Phosphate-3AMC7-Acyptene_FOPP	C31H40FN4O8P	c101	645.249	644.597	648.6808
Phosphate-3AMC7-D-HoCit_FOPP	C32H46FN6O9P	c087	707.297	706.6595	710.0528
Phosphate-3AMC7-bGln_FOPP	C30H41FN5O9P	c085	664.2548	663.237	666.9654
Phosphate-3AMC7-Gln_FOPP	C30H41FN5O9P	c102	664.2548	663.0556	668.3527
Phosphate-3AMC7-Cxcha_FOPP	C33H46FN4O8P	c088	675.2959	674.8134	677.8547
Phosphate-3AMC7-Valeram_FOPP	C32H43FN5O9P	c049	690.2704	689.3646	692.9398
Phosphate-3AMC7-LysFor_FOPP	C32H45FN5O9P	c108	692.2861	691.5269	695.3058
Phosphate-3AMC7-ArgMe2_FOPP	C33H49FN7O8P	c065	720.3286	719.7717	722.909
Phosphate-3AMC7-Ser_FOPP	C28H38FN4O9P	c112	623.2282	623.04	625.559
Phosphate-3AMC7-D-Ser_FOPPanchor	C28H38FN4O9P	c109	623.2282	623.0205	625.5002

Table 17. DPC Assembled Fragments LC/MS Analysis (Pool E)

Compound	Formula	Codon	Expected mass	Min. M/Z	Max. M/Z
Phosphate-3AMC7-3Pal_FOPP	C33FH41N5O8P	c109	684.2599	684.0737	686.4737
Phosphate-3AMC7-AFurBA_FOPP	C33FH42N4O9P	c088	687.2595	687.0734	689.4734
Phosphate-3AMC7-AMChxA_FOPP	C34FH48N4O8P	c078	689.3116	689.1254	691.5254
Phosphate-3AMC7-AThiBA_FOPP	C33FH42N4O8PS	c112	703.2367	703.0506	705.4506
Phosphate-3AMC7-AZPC_FOPP	C37FH45N5O9P	c065	752.2861	752.0999	754.4999
Phosphate-3AMC7-Ala2Fur_FOPP	C32FH40N4O9P	c049	673.2439	673.0577	675.4577
Phosphate-3AMC7-Ala4Thz_FOPP	C31FH39N5O8PS	c085	690.2163	690.0302	692.4302
Phosphate-3AMC7-CypAla_FOPP	C31FH42N4O8P	c102	647.2646	647.0785	649.4785
Phosphate-3AMC7-D_Ala4Thz_FOPP	C31FH39N5O8PS	c086	690.2163	690.0302	692.4302
Phosphate-3AMC7-Phe(MeO2)_FOPP	C36FH46N4O10P	c109	743.2857	743.0996	745.4996
Phosphate-3AMC7-Val_FOPP	C30FH42N4O8P	c101	635.2646	635.0785	637.4785
Phosphate-3AMC7-CNHoPhe_FOPP	C36FH43N5O8P	c108	722.2755	722.0894	724.4894

Table 18. DPC Assembled Fragments LC/MS Analysis (Pool H)

Compound	Formula	Codon	Expected mass	Min. M/Z	Max. M/Z
Phosphate-3AMC7-His(Bn)_FOPP	C38H46FN6O8P	c078	763.3021	762.6799	765.6587
Phosphate-3AMC7-ABztB_FOPP	C37H44FN4O8PS	c101	753.2523	752.5948	755.9703
Phosphate-3AMC7-F2HoPhe_FOPP	C35H42F3N4O8P	c088	733.2614	732.6077	736.3311
Phosphate-3AMC7-Freidam_FOPP	C35H49FN5O9P	c065	732.3174	731.4594	735.0368
Phosphate-3AMC7-Indanylgly_FOPP	C36H44FN4O8P	c102	709.2803	708.0261	712.7701
Phosphate-3AMC7-Styrylala_FOPP	C36H44FN4O8P	c108	709.2803	708.5441	712.3941
Phosphate-3AMC7-Phg_FOPP	C33H40FN4O8P	c085	669.249	668.5738	671.8665
Phosphate-3AMC7-Thi_FOPP	C32H40FN4O8PS	c087	689.221	688.5423	692.1912
Phosphate-3AMC7-Cha_FOPP	C34H48FN4O8P	c109	689.3116	688.6494	692.0565
Phosphate-3AMC7-Leu_FOPP	C31H44FN4O8P	c086	649.2803	648.3741	652.2355
Phosphate-3AMC7-DBip_FOPP	C40H46FN4O8P	c049	759.2959	758.7568	762.3244
Phosphate-3AMC7-LBip_FOPP	C40H46FN4O8P	c112	759.2959	758.4529	762.4684

Example 4 Selection of Anchor-based DPC Libraries

[0174] The ability of amino acid-based fragments to enhance the binding of an anchor when the two are conjugated to one another has been demonstrated. **FIG. 19** shows an example of binding of two 12-member anchor-based libraries to KDR. Two 12-member libraries containing the DNA conjugate of the anchor FOPP linked to a single diversity position, Pools H and A (see structures in **Tables 13 and 15**), were selected against the kinase protein target KDR.

[0175] Each member contains FOPP (see **FIG. 18**) as an anchor and a single amino acid as the conjugated variable fragment. Each individual member of each library is designated by codons 3a-3l. The relative binding of each member compared to the anchor control was determined at three different KDR concentrations as described in Methods (below). The binding of the corresponding linked DNA conjugate void of the anchor and the amino acid comprising the single point of diversity for each member was also determined (See text for discussion).

Methods

[0176] Indicated amounts of N-terminally 6x-His-tagged cytoplasmic domain from aa790-end (aa1357) of KDR (Upstate) was immobilized to Qiagen Ni^{2+}/nitrilotriacetic acid Superflow™ resin in 50 mM Tris, pH7.5, 300 mM NaCl, 270 mM sucrose, 0.03% Brij-35 for 2 hours at 4 °C. Using the same buffer, resin was washed three times and resuspended as a 25% slurry. 10 µL resin beds of KDR resins were pelleted and washed twice with 100 µL of binding buffer (25 mM Tris, pH 7.5/10 mM $MgCl_2$/1 mM Tris(2-carboxyethyl) phosphine /150 mM NaCl) and supernatants removed.

[0177] For binding experiments, 10 µL of the following mix was added to each resin: 12 nM 12-membered FOPP anchored libraries, 1 nM FOPP-DNA conjugate parent, 1 nM each of PTP1B inhibitors (see figure) conjugated to DNA, 5 µM decoy DNA (sequence = 5'CACTACAACACATCCCTCACCGTCAACACTCCATTCCCTCAC 3' (SEQ ID NO: 565), 25 mM Tris, pH7.5, 10 mM $MgCl_2$, 1 mM Tris(2-carboxyethyl) phosphine, 150 mM NaCl. For binding/competition experiments, 20 µL of the same mix at half the library and control concentrations in the presence of inhibitor and 0.5% DMSO was used. Libraries and resins were incubated at room temperature for one hour with slight agitation on a vortexer. 150 µL of binding buffer was added to each sample, resin resuspended and transferred separately to Ultrafree-MC 5 µm spin filter units (Millipore) and centrifuged briefly to remove buffer. Resin was washed with 2 X200 µL of binding buffer and recentrifuged. Resins were then resuspended in 100 µL binding buffer and transferred to 0.2 mL thin-walled PCR tube. Resins were centrifuged, supernatants removed and resins resuspended in 50 µL of 6 M guanidine-HCl. Resins were heated at 70 °C for twenty minutes, centrifuged, and supernatants transferred to 500 µL of PN buffer from Qiagen nucleotide removal kit. Samples were desalted according to manufacturer's protocol and eluted with 100 µL water.

[0178] Quantitative real-time PCR was used to quantitate small molecule-DNA conjugates in the applied material and the selected eluates. Briefly, the libraries and controls contain library-specific DNA sequences that can be used as a common 5' priming spot for each member of the library. The 3' primer is specific for the codons used to generate the DPC libraries. Biorad SYBRIQ was used to prepare mixes containing 0.5 µM 5' library-specific primer and 0.5 µM 3' codon-specific primers (one PCR reaction specific for each 3' codon). Five µL ($1/20^{th}$) of each sample was added to the PCR reaction mixes specific for each codon and quanitative real-time

PCR was performed on a Biorad ICycler. As a pre-binding control, the library mixes applied to the resins were diluted 1/100 and five μL of each were added to PCR mixes. Percent binding for each codon was determined by the relationship below and normalized to the anchor conjugate control.

$$\left[\frac{\text{the amount of PCR product in the binding sample}}{\text{the amount of PCR product in the pre-binding sample}}\right] \times 100$$

[0179] Relative to the anchor alone, the majority of conjugates in both libraries (Pools H and A) bound significantly more tightly. For example, in Pool H the conjugate containing the amino acid-based fragment, coded by codon 3f, bound approximately 10-fold more tightly than the anchor alone control. Likewise, for Pool A the conjugate containing the amino acid –based fragment coded by codon 3d bound 10-fold more tightly than the anchor alone control. As expected for the stringency of the selection process, the differential binding was most evident at the lower concentrations of the target protein. In control studies, the corresponding DNA-alone controls that did not contain any fragments or an anchor that could serve as a target binding element did not show significant binding to KDR relative to the anchor alone (note change in scale). These studies demonstrate the ability of fragments to enhance the binding of a known anchor.

Example 5 Discovery of Novel Ligands to Other Targets

[0180] Procedures of Example 4 may be applied to other targets of interest such as phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins. Examples of targets include BCR/Abl, BACE, HCV protease, P2Y(12), PTP1b, Renin, TNF-α and PAI-1.

[0181] The library of fragments may be selected against other targets such as BCR/Abl, using PCR to amplify sequences of binders. In one approach to the actual protein binding selections, DPC-fragment libraries are dissolved in aqueous binding buffer in one pot and equilibrated in the presence of immobilized target protein. Non-binders are washed away with buffer. Those molecules that may be binding through their attached DNA templates rather than through their fragment moieties are eliminated by washing the bound library with unfunctionalized DNA templates lacking PCR primer binding sites. Remaining ligands bound to the immobilized target are eluted.

[0182] To increase enrichment, one may iterate a selection by loading eluant from a first selection into a second selection to multiply the net enrichment. No intervening amplification of template is required. Iterating library selections can lead to very large enrichments of desired molecules. In certain embodiments, a first round of selection provides at least a 50-fold increase in the number of binding ligands. Preferably, the increase in enrichments is over 100-fold, more preferably over 1,000 fold, and even more preferably over 100,000-fold. Subsequent rounds of selection may further increase the enrichment 100-fold over the original library, preferably 1,000-fold, more preferably over 100,000-fold, and most preferably over 1,000,000-fold.

[0183] *In vitro* selections can also select for specificity in addition to binding affinity. Library screening methods for binding specificity typically require duplicating the entire screen for each target or non-target of interest.

[0184] In contrast, selections for specificity can be performed in a single experiment by selecting for target binding as well as for the inability to bind one or more non-targets. Thus, the library can be pre-depleted by removing library members that bind to a non-target. Alternatively, or in addition, selection for binding to the target molecule can be performed in the presence of an excess of one or more non-targets. To maximize specificity, the non-target can be a homologous molecule. If the target molecule is a protein, appropriate non-target proteins include, for example, a generally promiscuous protein such as an albumin. If the binding assay is designed to target only a specific portion of a target molecule, the non-target can be a variation on the molecule in which that portion has been changed or removed. See, e.g., U.S. Patent Application Publication No. 2004/0180412 A1 by Liu *et al.*

[0185] The DNA templates that encode and direct the syntheses of the target binding molecules may be amplified by any suitable technique, e.g., by PCR; nucleic acid sequence-based amplification (see, e.g., Compton, 1991, Nature, 350: 91-92), amplified anti-sense RNA (see, e.g., van Gelder *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85: 77652-77656); self-sustained sequence replication systems (Gnatelli *et al.*, 1990, Proc. Natl. Acad. Sci. USA 87: 18741878); polymerase-independent amplification (see, e.g., Schmidt *et al.*, 1997, Nucleic Acids Res. 25: 4797-4802, and in vivo amplification of plasmids carrying cloned DNA fragments. Description of PCR methods are found, for example, in Saiki *et al.*, 1985, Science 230: 1350-1354; Scharf *et al.*, 1986, Science 233: 1076-1078; and in U.S. Patent No. 4,683,202. Ligase-mediated amplification methods such as Ligase Chain Reaction (LCR) may also be used. In general, any

means allowing faithful, efficient amplification of selected nucleic acid sequences can be employed in the method of the present invention. It is preferable, although not necessary, that the proportionate representations of the sequences after amplification reflect the relative proportions of the sequences in the mixture before amplification.

[0186] Purification completes one cycle of translation, selection and amplification, yielding an enriched sub-population of DNA-fragments having binding affinities to the target protein.

[0187] The above process can be repeated until a subset of DPC-fragments are identified that bind to the target with desired affinity ranges, for example, "moderate affinity" (1 $\mu M < K_D <$ 10 μM), "moderately high affinity" (100 $nM < K_D < 1\ \mu M$), or "high affinity" ($K_D < 100$ nM, e.g., $K_D < 50$ nM or 20 nM, or "very high affinity" (1 nM or sub-nanomolar $< K_D < 10$ nM)). Additionally, deconvolution is performed on the set of binders from the mixture to obtain SAR of the target binding elements themselves. This allows one to infer where on a fragment substitution or other modifications may or may not be tolerated. Additionally, information can be obtained on SAR relating to the specific functionalities that should be tolerated in the subsequent DPC generated libraries for attaching fragments to each other or to other scaffolds.

[0188] To investigate a particular target binding element, the DNA sequence associated with the molecule can be sequenced using conventional approaches, which sequence can then be used to deconvolute the identity (e.g., structure and synthetic history) of the target binding element.

[0189] Sequencing can be performed by a standard dideoxy chain termination method, or by chemical sequencing, e.g., using the Maxam-Gilbert sequencing procedure. Alternatively, the sequence can be determined by hybridization to a chip. For example, a single-stranded DNA associated with a detectable moiety such as a fluorescent moiety is exposed to a chip bearing a large number of clonal populations of single-stranded nucleic acid analogs of known sequences, each clonal population being present at a particular addressable location on the chip. The unknown sequences are permitted to anneal to the chip sequences. The position of the detectable moieties on the chip then is determined. Based on the location of the detectable moiety and the immobilized sequence at that location, the sequence of the template can be determined. It is contemplated that large numbers of such oligonucleotides can be immobilized in an array on a chip or other solid support.

[0190] A combinatorial library can be prepared by a DPC process in which the identified target binding elements in the form of building blocks are incorporated. The target binding elements can be linked directly, via linking moieties or via scaffolds. The chemical assembly of the target binding elements using DPC to generate a library can be accomplished using chemical methodologies that have been established as amenable to DPC using strategies that have been shown appropriate for the multistep assembly of combinatorial libraries, as discussed above. This DPC-generated library is then selected against the target to identify those target binding elements that yield a more elaborated molecule with increased affinity for the target. See, e.g., U.S. Patent Application Publication No. 2004/0014090 A1 by Neri *et al.* and PCT International Publication No. WO 03/076943 A1; Gartner *et al.* Science, vol. 305, pp1601-1605, 2004; Doyon, *et al.*, JACS, vol. 125, pp 12372-12373, 2003.

[0191] The relative abundance of codons present in the library recovered from the selection is compared against the relative abundance of codons in the library prior to the selection. If a particular TBE, functionality, or scaffold, binds preferentially to the target, the relative abundance of the codons for the entity will increase as a result of the selection. If a particular entity is disfavored in binding, its relative frequency will decrease as a result of the selection. Additionally, optimal combinations of TBE's or functionalities, regardless of the scaffold in which they find themselves, may be preferred by the target binding site, and these interactions will be reflected in positive co-variance of pairs of codon frequencies. These data can be tabulated and analyzed to determine the optimal set of TBE's/codons to carry into a second or next round of selection.

[0192] The above example is envisioned to be applicable in general to other kinases, e.g., tyrosine kinases. Other exemplary kinases of therapeutic interest: VEGFR, PDGFR, EGFR, c-Kit, Flt-3, Src, Lck, Aurora, CDK's, JAK, IKK, p38, Raf, ERB B1&2, and JNK.

INCORPORATION BY REFERENCE

[0193] The entire disclosure of each of the publications and patent documents referred to herein is incorporated by reference in its entirety for all purposes to the same extent as if each individual publication or patent document were so individually denoted.

EQUIVALENTS

[0194] The invention may be embodied in other specific forms without departing form the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes that come within the meaning and range of equivalency of the claims are intended to be embraced therein.

WHAT IS CLAIMED IS:

CLAIMS

1. A method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule, wherein the target binding element has a K_D of 10 mM or lower, the method comprising:

(a) combining a target molecule with a plurality of pre-selected test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule, wherein each test molecule comprises a target binding element associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other target binding elements;

(b) harvesting a target binding element that binds to the target molecule with a K_D of 10 mM or lower; and

(c) determining the sequence of the oligonucleotide associated with the target binding element harvested in step (b) so as to identify the target binding element having a K_D of 10 mM or lower with the binding site.

2. The method of claim 1, wherein step (c) comprises:

amplifying the oligonucleotide associated with the target binding element harvested in step (b); and

determining the sequence of the amplified oligonucleotide so as to identify the target binding element having a K_D of 10 mM or lower with the binding site.

3. The method of claims 1 or 2, further comprising the step of, after step (a) but before step (b), washing away unbound target binding elements.

4. The method of claim 1, 2 or 3, further comprising the step of, before step (b), washing away target binding elements that bind to the target with K_D greater than 1 M.

5. The method of claim 1, wherein the target binding element has a mass ranging from 90 to 500 daltons.

6. The method of claim 5, wherein the target binding element has a mass ranging from 150 to 350 daltons.

7. The method of claim 1, wherein the target binding element has a K_D with the target molecule less than 1 nM.

8. The method of claim 1, wherein the target binding element has a K_D with the target molecule in the range from 1 nM to 100 nM.

9. The method of claim 1, wherein the target binding element has a K_D with the target molecule in the range from 100 nM to 10 μM.

10. The method of claim 1, wherein the target binding element has a K_D with the target molecule in the range from 10 μM to 100 μM.

11. The method of claim 1, wherein the target binding element has a K_D with the target molecule in the range from 100 μM to 10 mM.

12. The method of claim 1, wherein during step (c), the oligonucleotide is amplified by polymerase chain reaction.

13. The method of claim 11, wherein during step (c), a primer anneals to the amplification sequence.

14. An *in vitro* method for producing a molecule that binds to a preselected target molecule comprising a binding site, wherein the binding site comprises a first binding domain and a second binding domain, the method comprising the steps of:

(a) providing a template and a reagent, wherein

(i) the template comprises a first target binding element attached to a first oligonucleotide defining a first codon sequence, wherein the first target binding element has a first K_D with the first binding domain of the binding site, and

(ii) the reagent comprises a second target binding element attached to a second oligonucleotide defining a first anti-codon sequence capable of hybridizing to the codon sequence, wherein the second target binding element has a second K_D with the second binding domain; and

(b) combining the template and the reagent under conditions to permit the first codon sequence to hybridize to the first anti-codon sequence so as to bring the first and second target binding elements into reactive proximity whereupon the first and second target binding elements are chemically coupled to produce a reaction product that binds to the preselected target molecule.

15. The method of claim 14, wherein the reaction product has a K_D with the binding site less than

(i) the first K_D of the first target element with the first binding domain, and

(ii) the second K_D of the second target binding element with the second binding domain.

16. The method of claim 14 or 15 further comprising the step of:

(c) combining the reaction product with the target molecule to determine the binding characteristics of the reaction product.

17. The method of claim 15, wherein in step (a), the first K_D of the first target binding element with the first binding domain is sufficient to permit the first target binding element to bind to the first binding domain in the absence of the second target binding element.

18. The method of claim 15, wherein in step (a), the first K_D of the first target binding element with the first binding domain is insufficient to permit the first target binding element to bind to the first binding domain in the absence of the second target binding element.

19. The method of claim 17, wherein in step (a), the second K_D of the second target binding element with the second binding site is insufficient to permit the second target binding element to bind to the second binding domain in the absence of the first binding element.

20. The method of claim 14 or 15, wherein in step (a), the first target binding element is known to bind to the first binding domain of the binding site.

21. The method of claim 20, wherein the first target binding element is an anchor.

22. The method of claim 15, further comprising the step of selecting the reaction product.

23. The method of claim 15, wherein the codon identifies the first target binding element associated with the first oligonucleotide.

24. The method of claim 15, wherein the anti-codon identifies the second target binding element associated with the second oligonucleotide.

25. The method of claim 15, wherein the template comprises a plurality of different codons.

26. The method of claim 25, wherein a plurality of different reagents are combined with the template, and wherein each reagent comprises a different second target binding element attached to a corresponding, different oligonucleotide defining a corresponding anti-codon sequence, and wherein the anti-codon sequence is indicative of a particular second target binding element attached to the anti-codon sequence.

27. The method of claim 26, wherein the reaction product comprises a first target element coupled to a plurality of second target elements.

28. The method of any one of claims 14-27, further comprising the step of analyzing the sequence of the first oligonucleotide associated with the reaction product.

29. The method of claim 28, wherein the first oligonucleotide is analyzed by sequencing.

30. The method of any one of claims 1-29, wherein the sequence of the template is indicative of reaction product.

31. A composition comprising a plurality of test molecules, wherein each of substantially all of the test molecules comprises a target binding element associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other target binding elements.

32. The composition of claim 31, wherein at least some of the target binding elements has a K_D with a binding site greater than 10 μM.

33. The composition of claim 31, wherein substantially all of the target binding elements has a K_D with a binding site greater than 10 μM.

34. The composition of claim 31, wherein substantially all of the target binding elements has a molecular weight less than 400 daltons.

35. The composition of claim 31, wherein substantially all of the target binding elements are attached to the oligonucleotide via one or more functional groups associated with the target binding elements.

36. The composition of claim 35, wherein the functional group is selected from the group consisting of amines, carboxylic acids, acid chlorides, chloroformates, aldehydes, ketones, hydrazines, hydrazides, esters, sulphonyl chlorides, alcohols, phenols, azides, thiols, isocyanates, isothiocyanates, alkyl and aryl halides, epoxides, aziridines, enamines, acrylamides, enolethers, imidates, oximes, alkenes, and acetylenes.

37. The composition of claim 35, wherein the functional group is selected from the group consisting of amino groups, aniline groups, carboxylic groups and bifunctional groups having both an amine moiety and a carboxylic moiety.

38. The composition of claim 31, wherein at least some of the test molecules are not associated with an oligonucleotide.

39. The composition of claim 31, wherein each of substantially all of the target binding elements has a cLogP between -2 and 4.

40. The composition of claim 31, wherein each of substantially all of the target binding elements has 8 or fewer H-bond donors.

41. The composition of claim 40, wherein each of substantially all of the target binding elements has 4 or fewer H-bond acceptors.

42. The composition of claim 31, wherein each of substantially all of the target binding elements has 3 or fewer chiral centers.

43. The composition of claim 31, wherein each of substantially all of the target binding elements has 1 or more chiral center.

44. A composition comprising a plurality of test molecules, wherein each of at least some of the test molecules comprises two or more target binding elements and is associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

45. A composition comprising a plurality of test molecules, wherein each of substantially all of the test molecules comprises two or more target binding elements and is associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the two or more target binding elements, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other test molecules.

46. The composition of claim 44 or 45, wherein each of at least some of the target binding elements has a K_D of 10 mM or less with a binding site.

47. The composition of claim 44 or 45, wherein each of substantially all of the target binding elements has a K_D of 10 mM or less with a binding site.

48. The composition of claim 44 or 45, wherein for substantially all of the test molecules the product of the K_D's with a binding site of the corresponding two or more target binding elements associated with the oligonucleotide corresponding to a test molecule are 10 mM or less.

49. The composition of claim 44 or 45, wherein each of substantially all of the target binding elements has a molecular weight between 90 and 500 daltons.

50. The composition of claim 44 or 45, wherein for substantially all of the test molecules the sum of the molecular weight of the corresponding two or more target binding elements associated with the oligonucleotide corresponding to a test molecule is between 120 and 400 daltons.

51. The composition of claim 44 or 45, wherein each of substantially all of the target binding elements is linked to a functional group selected from the group consisting of primary amines, secondary amines, primary anilines, carboxylic acids and a bifunctional groups having both an amine moiety and an acid moiety.

52. A complex of a target molecule bound to a test molecule comprising two or more target binding elements, wherein the test molecule is associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the test molecule and (ii) contains an amplification sequence, wherein each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or more H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

53. A composition comprising a plurality of complexes wherein each complex comprises a target molecule bound to a test molecule comprising two or more target binding elements, wherein each test molecule is associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules, and wherein each of substantially all of the target binding elements comprises a functional group through which the target binding element is attached to the oligonucleotide.

54. A composition comprising a plurality of complexes wherein each complex comprises a target molecule bound to a test molecule comprising two or more target binding elements, wherein each test molecule is associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the test molecule, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences of other test molecules.

55. The composition of claim 54, wherein each of substantially all of the target binding elements comprises a functional group through which the target binding element is attached to the oligonucleotide.

56. The composition of claim 54, wherein each of substantially all of the target binding elements has a cLogP between -2 and 4.

57. The composition of claim 54, wherein each of substantially all of the target binding elements has 4 or fewer H-bond donors.

58. The composition of claim 54, wherein each of substantially all of the target binding elements has 8 or fewer H-bond acceptors.

59. The composition of claim 54, wherein each of substantially all the target binding elements has 3 or fewer chiral centers.

60. The composition of claim 54, wherein each of substantially all of the target binding elements has 1 or more chiral center.

61. A method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule, the method comprising:

(a) combining a target molecule with a plurality of test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule, wherein each test molecule comprises a target binding element associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules;

(b) harvesting a target binding element that binds to the target molecule binding site with a K_D of 10 mM or less;

(c) amplifying the oligonucleotide associated with the target binding element harvested in step (b); and

(d) determining the sequence of the amplified oligonucleotide so as to identify the target binding element having a K_D of 10 mM or less with a binding site,

wherein each of substantially all of the target binding elements has at least one of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

62. The method of claim 61, further comprising the step of, after step (a) but before step (b), washing away unbound target binding elements.

63. The method of claim 61 or 62, further comprising the step of, before step (b), washing away target binding elements that bind to the target with a K_D with a binding site greater than 1 M.

64. The method of claim 61, 62 or 63, wherein during step (c), the oligonucleotide is amplified by polymerase chain reaction.

65. The method of claim 64, wherein during step (c), a primer anneals to the amplification sequence.

66. The method of claim 65, wherein a polymerase extends the primer annealed to the amplification sequence.

67. A method for identifying a target binding element capable of binding to a binding domain disposed within a binding site of a target molecule, wherein the target binding element has a K_d of 10 mM or lower, the method comprising:

(a) combining a target molecule with a plurality of pre-selected test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule, wherein each test molecule comprises a target binding element associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequences associated with other target binding elements;

(b) harvesting a target binding element that binds to the target molecule with a K_d of 10 mM or lower; and

(c) determining the sequence of the oligonucleotide associated with the target binding element harvested in step (b) so as to identify the target binding element having a K_d of 10 mM or lower with the binding site.

68. An *in vitro* method for producing a molecule that binds to a preselected target molecule comprising a binding site, wherein the binding site comprises a first binding domain and a second binding domain, the method comprising the steps of:

(a) providing a template and a reagent, wherein

(i) the template comprises a first target binding element attached to a first oligonucleotide defining a first codon sequence, wherein the first target binding element has a first K_d with the first binding domain of the binding site, and

(ii) the reagent comprises a second target binding element attached to a second oligonucleotide defining a first anti-codon sequence capable of hybridizing to the codon sequence, wherein the second target binding element has a second K_d with the second binding domain; and

(b) combining the template and the reagent under conditions to permit the first codon sequence to hybridize to the first anti-codon sequence so as to bring the first and second target binding elements into reactive proximity whereupon the first and second target binding elements are chemically coupled to produce a reaction product that has a K_d with the binding site less than (i) the first K_d of the first target binding element with the first binding domain, and (ii) the second K_d of the second target binding element with the second binding domain.

69. A method for identifying a target binding element capable of binding to a target molecule, the method comprising:

(a) combining a target molecule with a plurality of test molecules under conditions that permit a test molecule to bind to a binding domain of the target molecule, wherein each test molecule comprises a target binding element associated with a corresponding oligonucleotide having a nucleotide sequence that (i) identifies the target binding element, (ii) contains an amplification sequence, and (iii) is substantially incapable of hybridizing to the nucleotide sequence associated with other test molecules;

(b) harvesting a target binding element that binds to the target molecule binding site with a K_d of 10 mM or less;

(c) amplifying the oligonucleotide associated with the target binding element harvested in step (b); and

(d) determining the sequence of the amplified oligonucleotide so as to identify the target binding element having a K_d of 10 mM or less with a binding site,

wherein each of substantially all of the target binding elements has all of the following characteristics: (i) a cLogP between -2 and 4, (ii) 4 or fewer H-bond donors, (iii) 8 or fewer H-bond acceptors, and (iv) a molecular weight between 90 and 500 daltons.

70. A method for identifying a compound having a desired binding affinity to a target molecule, the method comprising:

(a) providing a library comprising a plurality of test compounds, wherein each of the test compounds comprises (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound, wherein the common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule;

(b) providing a reference compound that comprises the common binding moiety;

(c) combining the target molecule, the library of test compounds, and the reference compound under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule;

(d) harvesting the test compounds that exhibit greater binding affinity to the target molecule than the reference compound; and

(e) determining the oligonucleotide sequences of the test compounds harvested thereby identifying the test compounds having a desired binding affinity to the target molecule.

71. The method of claim 70, wherein the test compounds are prepared by nucleic acid-templated synthesis.

72. The method of claim 70, wherein the bridging moiety is a part of the common binding moiety.

73. The method of claim 70, wherein the bridging moiety is a part of the scaffold binding moiety.

74. The method of claim 70, wherein the bridging moiety is a part of the common binding moiety and is a part of the scaffold binding moiety.

75. The method of claim 70, wherein the common binding moiety is a part of the scaffold moiety.

76. The method of claim 70, wherein the oligonucleotide is attached directly to the bridging moiety.

77. The method of claim 70, wherein the oligonucleotide is not attached directly to the bridging moiety.

78. The method of claim 70, wherein the oligonucleotide is attached directly to the scaffold moiety.

79. The method of claim 70, wherein the oligonucleotide is attached directly to the common binding moiety.

80. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain.

81. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain with a dissociation constant of 10 mM or lower.

82. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain with a dissociation constant of 100 μM or lower.

83. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain with a dissociation constant of 10 μM or lower.

84. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain with a dissociation constant of 1 μM or lower.

85. The method of claim 70, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a binding affinity to the second binding domain with a dissociation constant of 100 nM or lower.

86. The method of any of claims 70-85 wherein the target is selected from the group consisting of phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

87. A method for identifying a compound having a desired binding affinity to a target molecule, the method comprising:

(a) combining the target molecule, a plurality of test compounds, and a reference compound under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule, wherein (i) each of the plurality of test compounds comprises (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the associated test compound, (ii) the reference compound comprises the common binding moiety, and (iii) the common binding moiety has a binding affinity of 10 mM or lower to a first binding domain of the target molecule;

(b) determining the oligonucleotide sequences of the test compounds that bound to the target.

88. The method of claim 87, wherein the dissociation constant of the common binding moiety to the first binding domain of the target molecule is 100 μM or lower.

89. The method of claim 87, wherein the dissociation constant of the common binding moiety to the first binding domain of the target molecule is 10 μM or lower.

90. The method of claim 87, wherein the dissociation constant of the common binding moiety to the first binding domain of the target molecule is 100 nM or lower.

91. The method of claim 87, wherein the test compounds are prepared by nucleic acid-templated synthesis.

92. The method of claim 87, wherein the oligonucleotide is attached directly to the bridging moiety.

93. The method of claim 87, wherein the oligonucleotide is not attached directly to the bridging moiety.

94. The method of claim 87, wherein the oligonucleotide is attached directly to the scaffold moiety.

95. The method of claim 87, wherein the oligonucleotide is attached directly to the common binding moiety.

96. The method of claim 87, wherein the target molecule has a second binding domain and at least one of the harvested test compounds has a dissociation constant to the second binding domain of 10 mM or lower.

97. The method of any of claims 87-96 wherein the target is selected from the group consisting of kinases, phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

98. A method for detecting a second binding domain on a target molecule having a first binding domain, the method comprising:

(a) providing a test compound comprising (1) a first binding moiety having a binding affinity to the first binding domain of the target molecule, (2) a scaffold moiety connected to the first binding moiety through a bridging moiety, and (3) a defining oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the test compound, wherein the first binding moiety has a dissociation constant of 10 mM or lower to the first binding domain of the target molecule;

(b) determining the effect of the test compound on the binding of a reference compound to the target molecule, wherein the reference compound comprises the first binding moiety; and

(c) analyzing the data collecting in (b) to detect the presence of a second binding domain on the target molecule.

99. The method of claim 98 further comprising

(d) determining the binding affinity of the scaffold moiety of the test molecule to the second binding domain of the target molecule.

100. The method of claim 98, wherein the test compounds are prepared by nucleic acid-templated synthesis.

101. The method of claim 98, wherein the bridging moiety is a part of the common binding moiety.

102. The method of claim 98, wherein the bridging moiety is a part of the scaffold binding moiety.

103. The method of claim 98, wherein the bridging moiety is a part of the common binding moiety and is a part of the scaffold binding moiety.

104. The method of claim 98, wherein the common binding moiety is a part of the scaffold moiety.

105. The method of claim 98, wherein the oligonucleotide is attached directly to the bridging moiety.

106. The method of claim 98, wherein the oligonucleotide is not attached directly to the bridging moiety.

107. The method of claim 98, wherein the oligonucleotide is attached directly to the scaffold moiety.

108. The method of claim 98, wherein the oligonucleotide is attached directly to the first binding moiety.

109. The method of claim 98, wherein the target molecule has a second binding domain and at least one of the test compounds has a dissociation constant to the second binding domain of 10 mM or lower.

110. The method of claim 98, wherein the target molecule has a second binding domain and at least one of the test compounds has a dissociation constant to the second binding domain of 100 μM or lower.

111. The method of any of claims 98-109 wherein the target is selected from the group consisting of kinases, phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

112. A method for identifying a compound having a desired binding affinity to a target molecule, the method comprising:

(a) providing a library comprising a plurality of test compounds, wherein each of the test compound comprises (1) a common binding moiety, (2) a scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound, wherein the common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule;

(b) combining the target molecule and the plurality of test compound under conditions that permit binding of one or more of the plurality of test compounds to the target molecule if such test compounds with desired binding affinity are present;

(c) harvesting the test compounds bound to the target; and

(d) determining the oligonucleotide sequences of the test compounds harvested thereby identifying the test compounds having a desired binding affinity to the target molecule.

113. The method of claim 112, wherein the test compounds are prepared by nucleic acid-templated synthesis.

114. The method of claim 112, wherein the bridging moiety is a part of the common binding moiety.

115. The method of claim 112, wherein the bridging moiety is a part of the scaffold moiety.

116. The method of claim 112, wherein the bridging moiety is a part of the common binding moiety and is a part of the scaffold moiety.

117. The method of claim 112, wherein the common binding moiety is a part of the scaffold moiety.

118. The method of claim 112, wherein the oligonucleotide is attached directly to the bridging moiety.

119. The method of claim 112, wherein the oligonucleotide is not attached directly to the bridging moiety.

120. The method of claim 112, wherein the oligonucleotide is attached directly to the scaffold moiety.

121. The method of claim 112, wherein the oligonucleotide is attached directly to the common binding moiety.

122. The method of claim 112, wherein the target molecule has a second binding domain and at least one of the test compounds has a dissociation constant to the second binding domain of 10 mM or lower

123. The method of any of claims 112-122 wherein the target is selected from the group consisting of kinases, phosphatases, proteases, receptors, ion channels, oxidases and reductases, catabolic and anabolic enzymes, pumps, and electron transport proteins.

124. A method for identifying a compound having a desired binding affinity to a target molecule, the method comprising:

(a) providing a library comprising two subsets of test compounds, wherein each of the first subset of test compounds comprises (1) a common binding moiety, (2) a first scaffold moiety connected to the common binding moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound, wherein the common binding moiety has a dissociation constant of 10 mM or lower to a first binding domain of the target molecule, and wherein each of the second subset of test compounds comprises (1) a second scaffold moiety, and (2) an oligonucleotide having a nucleotide sequence informative of the structural or synthetic information of the associated test compound;

(b) providing a reference compound that comprises the common binding moiety;

(c) combining the target molecule, the library comprising the two subsets of test compounds, and the reference compound under conditions that permit the plurality of test compounds and the reference compound to compete for binding to the target molecule;

(d) harvesting the test compounds that exhibit greater binding affinity to the target molecule than the reference compound; and

(e) determining the oligonucleotide sequences of the test compounds harvested thereby identifying the test compounds having a desired binding affinity to the target molecule.

125. The method of claim 124 wherein the first scaffold and the second scaffold are the same scaffold.

126. A compound identified by any of methods of claims 70-97 and 112-125.

127. A library of chemical compounds, the library comprising a plurality of compounds, wherein each of the compounds comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety, wherein (i) the compounds are prepared by one or more nucleic-acid-templated chemical reactions and (ii) the first moiety has a dissociation constant of 10 mM or lower to a binding domain of the target molecule.

128. A compound that comprises (1) a first moiety, (2) a second moiety connected to the first moiety through a bridging moiety, and (3) an oligonucleotide having a nucleotide sequence informative of the structure or synthetic information of the second moiety, wherein (i) the compounds are prepared by one or more nucleic-acid-templated chemical reactions and (ii) the first moiety has a dissociation constant of 10 mM or lower to a binding domain of the target molecule.

FIG. 1

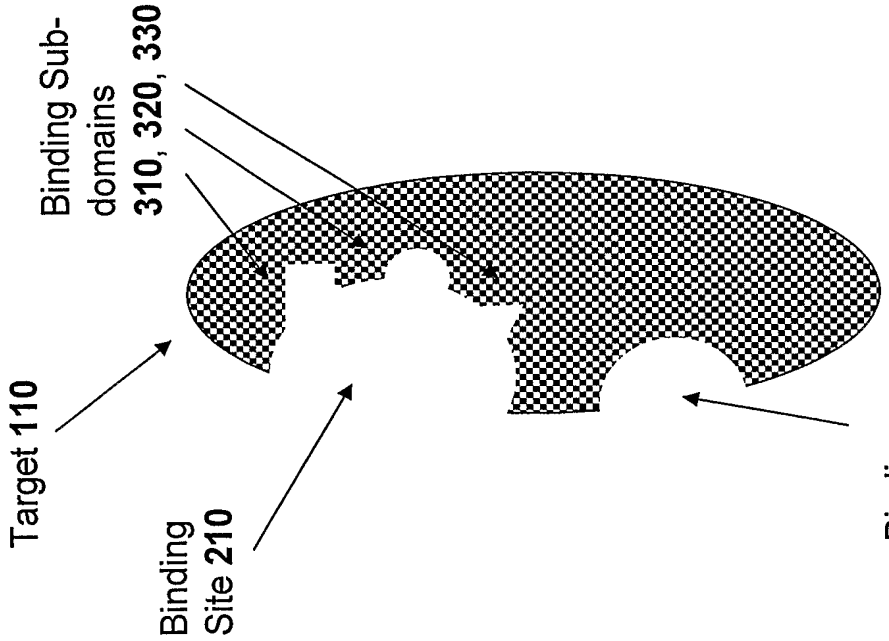

Target 110
Binding Sub-domains 310, 320, 330
Binding Site 210
Binding Site 220

FIG.2

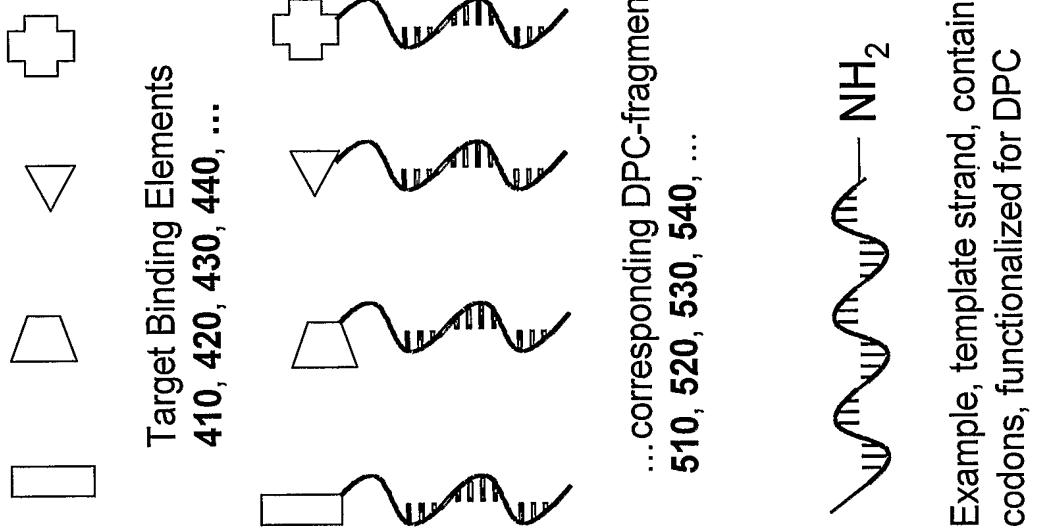

Target Binding Elements 410, 420, 430, 440, ...
Reagent strands (containing anti-codons), or detection strands
...corresponding DPC-fragments 510, 520, 530, 540, ...
NH_2
Example, template strand, containing codons, functionalized for DPC

FIG. 3

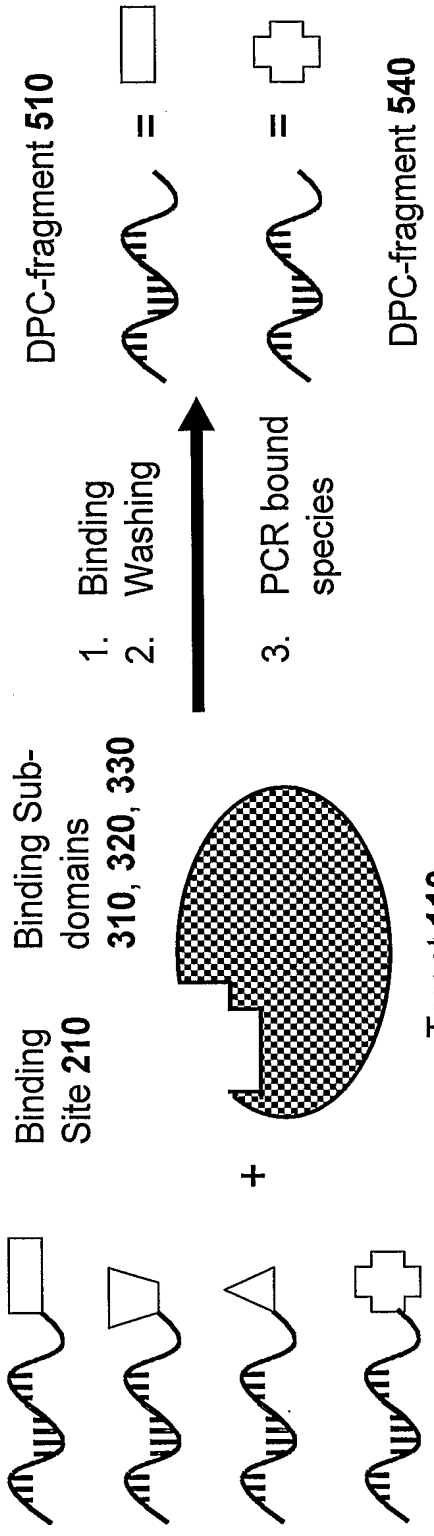

DPC-fragments
510, 520, 530, 540, ...
+
Binding Site 210
Binding Sub-domains 310, 320, 330
Target 110
1. Binding
2. Washing
3. PCR bound species
DPC-fragment 510
DPC-fragment 540

FIG. 4

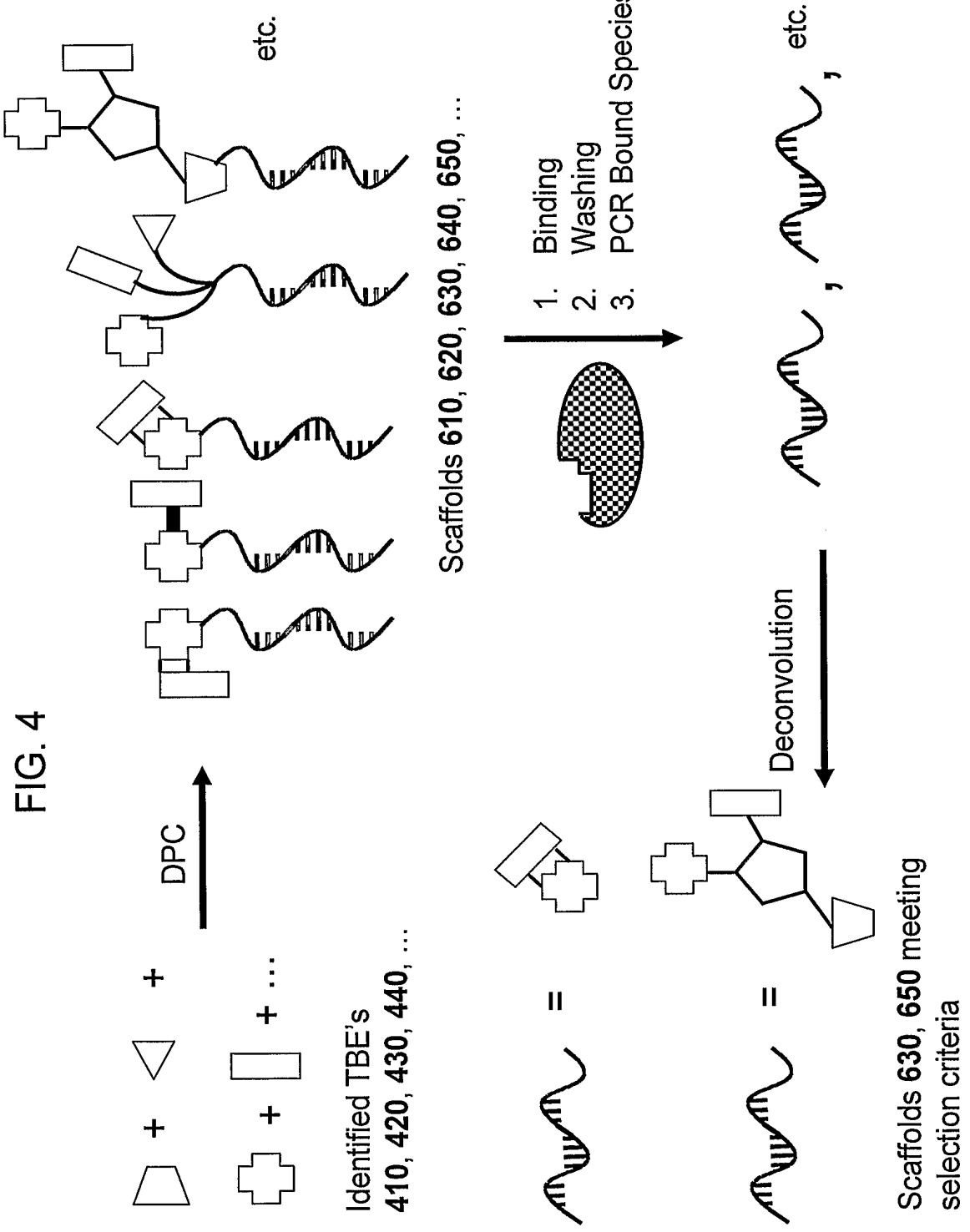
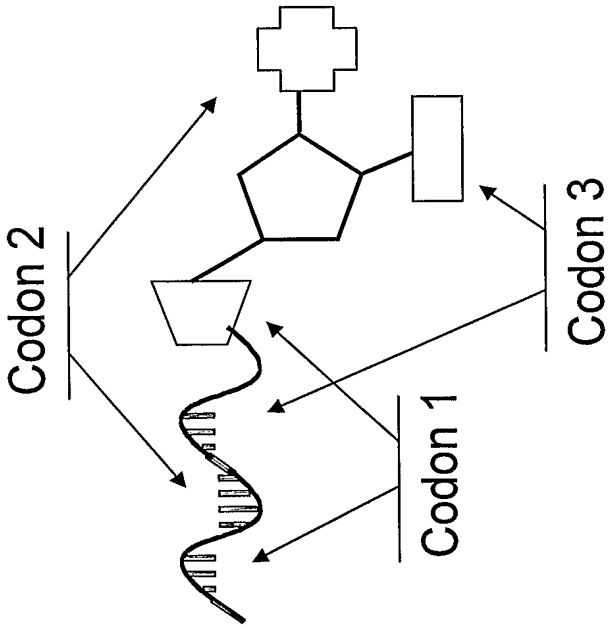

DPC
Identified TBE's
410, 420, 430, 440, ...
Scaffolds 610, 620, 630, 640, 650, ...
1. Binding
2. Washing
3. PCR Bound Species
Deconvolution
Scaffolds 630, 650 meeting
selection criteria
etc.
etc.

FIG. 5



- TBE's identified by enriched/depleted codons
- Reagents and codon may be re-used, available for iteration
- TBE interactions revealed through co-variance of codons

FIG. 6

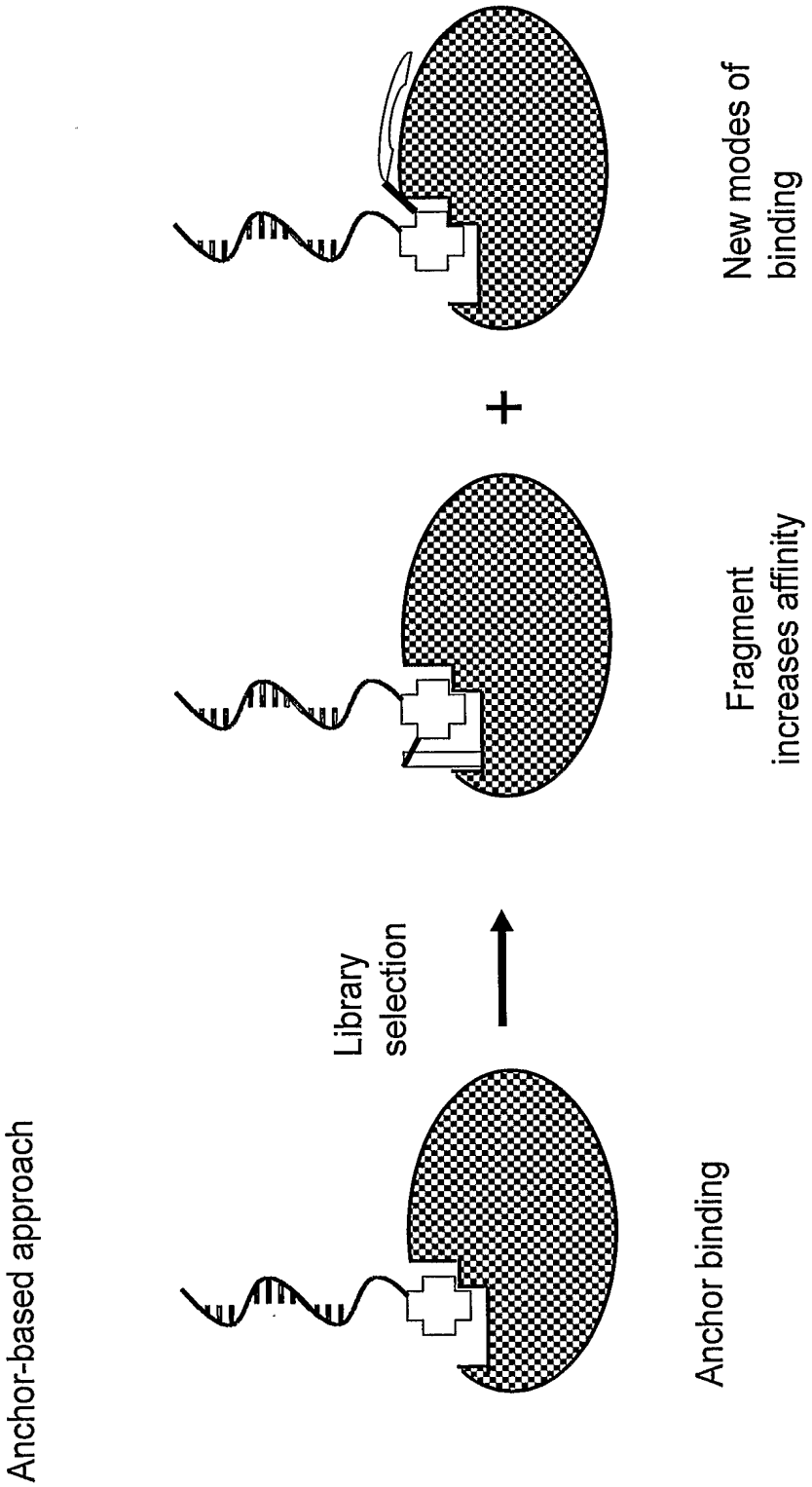

Anchor-based approach
Library selection
Anchor binding
Fragment increases affinity
New modes of binding

FIG. 7

(All sequences below presented in 5' to 3' direction)
GCTTGTCTACACACACACCTGGAG (SEQ ID NO: 1)
GCTTGTCTACACACATACACCTGGAG (SEQ ID NO: 2)
GCTTGTCTACACACCAACACCTGGAG (SEQ ID NO: 3)
GCTTGTCTACACACCCACACCTGGAG (SEQ ID NO: 4)
GCTTGTCTACACACCTACACCTGGAG (SEQ ID NO: 5)
GCTTGTCTACACACGAACACCTGGAG (SEQ ID NO: 6)
GCTTGTCTACACACGCACACCTGGAG (SEQ ID NO: 7)
GCTTGTCTACACACTAACACCTGGAG (SEQ ID NO: 8)
GCTTGTCTACACACTCACACCTGGAG (SEQ ID NO: 9)
GCTTGTCTACACACTTACACCTGGAG (SEQ ID NO: 10)
GCTTGTCTACACATACACACCTGGAG (SEQ ID NO: 11)
GCTTGTCTACACATCAACACCTGGAG (SEQ ID NO: 12)
GCTTGTCTACACATCTACACCTGGAG (SEQ ID NO: 13)
GCTTGTCTACACCAAAACACCTGGAG (SEQ ID NO: 14)
GCTTGTCTACACCAACACACCTGGAG (SEQ ID NO: 15)
GCTTGTCTACACCAATACACCTGGAG (SEQ ID NO: 16)
GCTTGTCTACACCACCACACCTGGAG (SEQ ID NO: 17)
GCTTGTCTACACCACTACACCTGGAG (SEQ ID NO: 18)
GCTTGTCTACACCATAACACCTGGAG (SEQ ID NO: 19)
GCTTGTCTACACCATCACACCTGGAG (SEQ ID NO: 20)
GCTTGTCTACACCATTACACCTGGAG (SEQ ID NO: 21)
GCTTGTCTACACCCAAACACCTGGAG (SEQ ID NO: 22)
GCTTGTCTACACCCACACACCTGGAG (SEQ ID NO: 23)
GCTTGTCTACACCCATACACCTGGAG (SEQ ID NO: 24)
GCTTGTCTACACCCCAACACCTGGAG (SEQ ID NO: 25)
GCTTGTCTACACCCCTACACCTGGAG (SEQ ID NO: 26)
GCTTGTCTACACCCGAACACCTGGAG (SEQ ID NO: 27)
GCTTGTCTACACCCTAACACCTGGAG (SEQ ID NO: 28)
GCTTGTCTACACCCTCACACCTGGAG (SEQ ID NO: 29)
GCTTGTCTACACCCTTACACCTGGAG (SEQ ID NO: 30)
GCTTGTCTACACCGAAACACCTGGAG (SEQ ID NO: 31)
GCTTGTCTACACCGCAACACCTGGAG (SEQ ID NO: 32)
GCTTGTCTACACCGCTACACCTGGAG (SEQ ID NO: 33)
GCTTGTCTACACCTAAACACCTGGAG (SEQ ID NO: 34)
GCTTGTCTACACCTACACACCTGGAG (SEQ ID NO: 35)
GCTTGTCTACACCTCAACACCTGGAG (SEQ ID NO: 36)
GCTTGTCTACACCTCTACACCTGGAG (SEQ ID NO: 37)
GCTTGTCTACACCTTCACACCTGGAG (SEQ ID NO: 38)
GCTTGTCTACACGAAAACACCTGGAG (SEQ ID NO: 39)
GCTTGTCTACACGAACACACCTGGAG (SEQ ID NO: 40)
GCTTGTCTACACGAATACACCTGGAG (SEQ ID NO: 41)
GCTTGTCTACACGACCACACCTGGAG (SEQ ID NO: 42)
GCTTGTCTACACGACTACACCTGGAG (SEQ ID NO: 43)
GCTTGTCTACACGCAAACACCTGGAG (SEQ ID NO: 44)
GCTTGTCTACACGCACACACCTGGAG (SEQ ID NO: 45)
GCTTGTCTACACGCATACACCTGGAG (SEQ ID NO: 46)

FIG. 7 (cont'd)

GCTTGTCTACACGCCAACACCTGGAG (SEQ ID NO: 47)
GCTTGTCTACACGCCTACACCTGGAG (SEQ ID NO: 48)
GCTTGTCTACACGCTAACACCTGGAG (SEQ ID NO: 49)
GCTTGTCTACACGCTCACACCTGGAG (SEQ ID NO: 50)
GCTTGTCTACACGCTTACACCTGGAG (SEQ ID NO: 51)
GCTTGTCTACACGGAAACACCTGGAG (SEQ ID NO: 52)
GCTTGTCTACACGGCTACACCTGGAG (SEQ ID NO: 53)
GCTTGTCTACACTAACACACCTGGAG (SEQ ID NO: 54)
GCTTGTCTACACTACCACACCTGGAG (SEQ ID NO: 55)
GCTTGTCTACACTACTACACCTGGAG (SEQ ID NO: 56)
GCTTGTCTACACTATCACACCTGGAG (SEQ ID NO: 57)
GCTTGTCTACACTCAAACACCTGGAG (SEQ ID NO: 58)
GCTTGTCTACACTCACACACCTGGAG (SEQ ID NO: 59)
GCTTGTCTACACTCATACACCTGGAG (SEQ ID NO: 60)
GCTTGTCTACACTCGCACACCTGGAG (SEQ ID NO: 61)
GCTTGTCTACACTCTAACACCTGGAG (SEQ ID NO: 62)
GCTTGTCTACACTCTCACACCTGGAG (SEQ ID NO: 63)
GCTTGTCTACACTCTTACACCTGGAG (SEQ ID NO: 64)
GCTTGTCTACACTGCCACACCTGGAG (SEQ ID NO: 65)
GCTTGTCTACACTGCTACACCTGGAG (SEQ ID NO: 66)
GCTTGTCTACACTGGAACACCTGGAG (SEQ ID NO: 67)
GCTTGTCTACACTGTCACACCTGGAG (SEQ ID NO: 68)
GCTTGTCTACACTTACACACCTGGAG (SEQ ID NO: 69)
GCTTGTCTACACTTCAACACCTGGAG (SEQ ID NO: 70)
GCTTGTCTACACTTCTACACCTGGAG (SEQ ID NO: 71)
GCTTGTCTACATAACCACACCTGGAG (SEQ ID NO: 72)
GCTTGTCTACATACACACACCTGGAG (SEQ ID NO: 73)
GCTTGTCTACATACCAACACCTGGAG (SEQ ID NO: 74)
GCTTGTCTACATACCCACACCTGGAG (SEQ ID NO: 75)
GCTTGTCTACATACCTACACCTGGAG (SEQ ID NO: 76)
GCTTGTCTACATACGAACACCTGGAG (SEQ ID NO: 77)
GCTTGTCTACATACGCACACCTGGAG (SEQ ID NO: 78)
GCTTGTCTACATACTCACACCTGGAG (SEQ ID NO: 79)
GCTTGTCTACATCAACACACCTGGAG (SEQ ID NO: 80)
GCTTGTCTACATCACCACACCTGGAG (SEQ ID NO: 81)
GCTTGTCTACATCACTACACCTGGAG (SEQ ID NO: 82)
GCTTGTCTACATCATCACACCTGGAG (SEQ ID NO: 83)
GCTTGTCTACATCCCAACACCTGGAG (SEQ ID NO: 84)
GCTTGTCTACATCCCCACACCTGGAG (SEQ ID NO: 85)
GCTTGTCTACATCCCTACACCTGGAG (SEQ ID NO: 86)
GCTTGTCTACATCCTAACACCTGGAG (SEQ ID NO: 87)
GCTTGTCTACATCCTCACACCTGGAG (SEQ ID NO: 88)
GCTTGTCTACATCCTTACACCTGGAG (SEQ ID NO: 89)
GCTTGTCTACATCGCAACACCTGGAG (SEQ ID NO: 90)
GCTTGTCTACATCGCCACACCTGGAG (SEQ ID NO: 91)
GCTTGTCTACATCGCTACACCTGGAG (SEQ ID NO: 92)
GCTTGTCTACATCTACACACCTGGAG (SEQ ID NO: 93)

FIG. 7 (cont'd)

GCTTGTCTACATCTCAACACCTGGAG (SEQ ID NO: 94)
GCTTGTCTACATCTCTACACCTGGAG (SEQ ID NO: 95)
GCTTGTCTACATCTTCACACCTGGAG (SEQ ID NO: 96)
GCTTGTCTACATTACCACACCTGGAG (SEQ ID NO: 97)
GCTTGTCTACCAAAACACACCTGGAG (SEQ ID NO: 98)
GCTTGTCTACCAAACCACACCTGGAG (SEQ ID NO: 99)
GCTTGTCTACCAAACTACACCTGGAG (SEQ ID NO: 100)
GCTTGTCTACCAAATCACACCTGGAG (SEQ ID NO: 101)
GCTTGTCTACCAACACACACCTGGAG (SEQ ID NO: 102
GCTTGTCTACCAACATACACCTGGAG (SEQ ID NO: 103)
GCTTGTCTACCAACCAACACCTGGAG (SEQ ID NO: 104)
GCTTGTCTACCAACCCACACCTGGAG (SEQ ID NO: 105)
GCTTGTCTACCAACCTACACCTGGAG (SEQ ID NO: 106)
GCTTGTCTACCAACGAACACCTGGAG (SEQ ID NO: 107)
GCTTGTCTACCAACGCACACCTGGAG (SEQ ID NO: 108)
GCTTGTCTACCAACTAACACCTGGAG (SEQ ID NO: 109)
GCTTGTCTACCAACTCACACCTGGAG (SEQ ID NO: 110)
GCTTGTCTACCAACTTACACCTGGAG (SEQ ID NO: 111)
GCTTGTCTACCAATACACACCTGGAG (SEQ ID NO: 112)
GCTTGTCTACCAATCAACACCTGGAG (SEQ ID NO: 113)
GCTTGTCTACCAATCTACACCTGGAG (SEQ ID NO: 114)
GCTTGTCTACCACACCACACCTGGAG (SEQ ID NO: 115)
GCTTGTCTACCACACTACACCTGGAG (SEQ ID NO: 116)
GCTTGTCTACCACATAACACCTGGAG (SEQ ID NO: 117)
GCTTGTCTACCACATCACACCTGGAG (SEQ ID NO: 118)
GCTTGTCTACCACATTACACCTGGAG (SEQ ID NO: 119)
GCTTGTCTACCACCAAACACCTGGAG (SEQ ID NO: 120)
GCTTGTCTACCACCACACACCTGGAG (SEQ ID NO: 121)
GCTTGTCTACCACCATACACCTGGAG (SEQ ID NO: 122)
GCTTGTCTACCACCCAACACCTGGAG (SEQ ID NO: 123)
GCTTGTCTACCACCCTACACCTGGAG (SEQ ID NO: 124)
GCTTGTCTACCACCGAACACCTGGAG (SEQ ID NO: 125)
GCTTGTCTACCACCTAACACCTGGAG (SEQ ID NO: 126)
GCTTGTCTACCACCTCACACCTGGAG (SEQ ID NO: 127)
GCTTGTCTACCACCTTACACCTGGAG (SEQ ID NO: 128)
GCTTGTCTACCACGAAACACCTGGAG (SEQ ID NO: 129)
GCTTGTCTACCACGCAACACCTGGAG (SEQ ID NO: 130)
GCTTGTCTACCACGCTACACCTGGAG (SEQ ID NO: 131)
GCTTGTCTACCACGGAACACCTGGAG (SEQ ID NO: 132)
GCTTGTCTACCACTAAACACCTGGAG (SEQ ID NO: 133)
GCTTGTCTACCACTACACACCTGGAG (SEQ ID NO: 134)
GCTTGTCTACCACTCAACACCTGGAG (SEQ ID NO: 135)
GCTTGTCTACCACTCTACACCTGGAG (SEQ ID NO: 136)
GCTTGTCTACCACTTCACACCTGGAG (SEQ ID NO: 137)
GCTTGTCTACCATAACACACCTGGAG (SEQ ID NO: 138)
GCTTGTCTACCATACCACACCTGGAG (SEQ ID NO: 139)
GCTTGTCTACCATACTACACCTGGAG (SEQ ID NO: 140)

FIG. 7 (cont'd)

GCTTGTCTACCATCAAACACCTGGAG (SEQ ID NO: 141)
GCTTGTCTACCATCACACACCTGGAG (SEQ ID NO: 142)
GCTTGTCTACCATCATACACCTGGAG (SEQ ID NO: 143)
GCTTGTCTACCATCCCACACCTGGAG (SEQ ID NO: 144)
GCTTGTCTACCATCCTACACCTGGAG (SEQ ID NO: 145)
GCTTGTCTACCATCGCACACCTGGAG (SEQ ID NO: 146)
GCTTGTCTACCATCTAACACCTGGAG (SEQ ID NO: 147)
GCTTGTCTACCATCTCACACCTGGAG (SEQ ID NO: 148)
GCTTGTCTACCATCTTACACCTGGAG (SEQ ID NO: 149)
GCTTGTCTACCATTACACACCTGGAG (SEQ ID NO: 150)
GCTTGTCTACCCAAAAACACCTGGAG (SEQ ID NO: 151)
GCTTGTCTACCCAAACACACCTGGAG (SEQ ID NO: 152)
GCTTGTCTACCCAAATACACCTGGAG (SEQ ID NO: 153)
GCTTGTCTACCCAACCACACCTGGAG (SEQ ID NO: 154)
GCTTGTCTACCCAACTACACCTGGAG (SEQ ID NO: 155)
GCTTGTCTACCCAATAACACCTGGAG (SEQ ID NO: 156)
GCTTGTCTACCCAATCACACCTGGAG (SEQ ID NO: 157)
GCTTGTCTACCCACACACACCTGGAG (SEQ ID NO: 158)
GCTTGTCTACCCACATACACCTGGAG (SEQ ID NO: 159)
GCTTGTCTACCCACCAACACCTGGAG (SEQ ID NO: 160)
GCTTGTCTACCCACCTACACCTGGAG (SEQ ID NO: 161)
GCTTGTCTACCCACGAACACCTGGAG (SEQ ID NO: 162)
GCTTGTCTACCCACTAACACCTGGAG (SEQ ID NO: 163)
GCTTGTCTACCCACTCACACCTGGAG (SEQ ID NO: 164)
GCTTGTCTACCCACTTACACCTGGAG (SEQ ID NO: 165)
GCTTGTCTACCCATAAACACCTGGAG (SEQ ID NO: 166)
GCTTGTCTACCCATACACACCTGGAG (SEQ ID NO: 167)
GCTTGTCTACCCATCAACACCTGGAG (SEQ ID NO: 168)
GCTTGTCTACCCATCTACACCTGGAG (SEQ ID NO: 169)
GCTTGTCTACCCCAAAACACCTGGAG (SEQ ID NO: 170)
GCTTGTCTACCCCAACACACCTGGAG (SEQ ID NO: 171)
GCTTGTCTACCCCAATACACCTGGAG (SEQ ID NO: 172)
GCTTGTCTACCCCACTACACCTGGAG (SEQ ID NO: 173)
GCTTGTCTACCCCATAACACCTGGAG (SEQ ID NO: 174)
GCTTGTCTACCCCATCACACCTGGAG (SEQ ID NO: 175)
GCTTGTCTACCCCATTACACCTGGAG (SEQ ID NO: 176)
GCTTGTCTACCCCCAAACACCTGGAG (SEQ ID NO: 177)
GCTTGTCTACCCCCATACACCTGGAG (SEQ ID NO: 178)
GCTTGTCTACCCCCTAACACCTGGAG (SEQ ID NO: 179)
GCTTGTCTACCCCCTTACACCTGGAG (SEQ ID NO: 180)
GCTTGTCTACCCCGAAACACCTGGAG (SEQ ID NO: 181)
GCTTGTCTACCCCTAAACACCTGGAG (SEQ ID NO: 182)
GCTTGTCTACCCCTACACACCTGGAG (SEQ ID NO: 183)
GCTTGTCTACCCCTCAACACCTGGAG (SEQ ID NO: 184)
GCTTGTCTACCCCTCTACACCTGGAG (SEQ ID NO: 185)
GCTTGTCTACCCCTTCACACCTGGAG (SEQ ID NO: 186)
GCTTGTCTACCCGAAAACACCTGGAG (SEQ ID NO: 187)

FIG. 7 (cont'd)

GCTTGTCTACCCGAACACACCTGGAG (SEQ ID NO: 188)
GCTTGTCTACCCGAATACACCTGGAG (SEQ ID NO: 189)
GCTTGTCTACCCGACTACACCTGGAG (SEQ ID NO: 190)
GCTTGTCTACCCGCAAACACCTGGAG (SEQ ID NO: 191)
GCTTGTCTACCCGCATACACCTGGAG (SEQ ID NO: 192)
GCTTGTCTACCCGCTAACACCTGGAG (SEQ ID NO: 193)
GCTTGTCTACCCGCTTACACCTGGAG (SEQ ID NO: 194)
GCTTGTCTACCCTAAAACACCTGGAG (SEQ ID NO: 195)
GCTTGTCTACCCTAACACACCTGGAG (SEQ ID NO: 196)
GCTTGTCTACCCTACCACACCTGGAG (SEQ ID NO: 197)
GCTTGTCTACCCTACTACACCTGGAG (SEQ ID NO: 198)
GCTTGTCTACCCTATCACACCTGGAG (SEQ ID NO: 199)
GCTTGTCTACCCTCAAACACCTGGAG (SEQ ID NO: 200)
GCTTGTCTACCCTCACACACCTGGAG (SEQ ID NO: 201)
GCTTGTCTACCCTCATACACCTGGAG (SEQ ID NO: 202)
GCTTGTCTACCCTCTAACACCTGGAG (SEQ ID NO: 203)
GCTTGTCTACCCTCTCACACCTGGAG (SEQ ID NO: 204)
GCTTGTCTACCCTCTTACACCTGGAG (SEQ ID NO: 205)
GCTTGTCTACCCTGCTACACCTGGAG (SEQ ID NO: 206)
GCTTGTCTACCCTGGAACACCTGGAG (SEQ ID NO: 207)
GCTTGTCTACCCTGTCACACCTGGAG (SEQ ID NO: 208)
GCTTGTCTACCCTTACACACCTGGAG (SEQ ID NO: 209)
GCTTGTCTACCCTTCAACACCTGGAG (SEQ ID NO: 210)
GCTTGTCTACCCTTCTACACCTGGAG (SEQ ID NO: 211)
GCTTGTCTACCGAAAAACACCTGGAG (SEQ ID NO: 212)
GCTTGTCTACCGAAACACACCTGGAG (SEQ ID NO: 213)
GCTTGTCTACCGAAATACACCTGGAG (SEQ ID NO: 214)
GCTTGTCTACCGAACCACACCTGGAG (SEQ ID NO: 215)
GCTTGTCTACCGAACTACACCTGGAG (SEQ ID NO: 216)
GCTTGTCTACCGAATAACACCTGGAG (SEQ ID NO: 217)
GCTTGTCTACCGAATCACACCTGGAG (SEQ ID NO: 218)
GCTTGTCTACCGACCAACACCTGGAG (SEQ ID NO: 219)
GCTTGTCTACCGACCTACACCTGGAG (SEQ ID NO: 220)
GCTTGTCTACCGACGAACACCTGGAG (SEQ ID NO: 221)
GCTTGTCTACCGACTAACACCTGGAG (SEQ ID NO: 222)
GCTTGTCTACCGACTCACACCTGGAG (SEQ ID NO: 223)
GCTTGTCTACCGACTTACACCTGGAG (SEQ ID NO: 224)
GCTTGTCTACCGCAAAACACCTGGAG (SEQ ID NO: 225)
GCTTGTCTACCGCAACACACCTGGAG (SEQ ID NO: 226)
GCTTGTCTACCGCAATACACCTGGAG (SEQ ID NO: 227)
GCTTGTCTACCGCACTACACCTGGAG (SEQ ID NO: 228)
GCTTGTCTACCGCATAACACCTGGAG (SEQ ID NO: 229)
GCTTGTCTACCGCATCACACCTGGAG (SEQ ID NO: 230)
GCTTGTCTACCGCATTACACCTGGAG (SEQ ID NO: 231)
GCTTGTCTACCGCCAAACACCTGGAG (SEQ ID NO: 232)
GCTTGTCTACCGCCATACACCTGGAG (SEQ ID NO: 233)
GCTTGTCTACCGCCTAACACCTGGAG (SEQ ID NO: 234)

FIG. 7 (cont'd)

GCTTGTCTACCGCCTTACACCTGGAG (SEQ ID NO: 235)
GCTTGTCTACCGCTAAACACCTGGAG (SEQ ID NO: 236)
GCTTGTCTACCGCTACACACCTGGAG (SEQ ID NO: 237)
GCTTGTCTACCGCTCAACACCTGGAG (SEQ ID NO: 238)
GCTTGTCTACCGCTCTACACCTGGAG (SEQ ID NO: 239)
GCTTGTCTACCGCTTCACACCTGGAG (SEQ ID NO: 240)
GCTTGTCTACCTAAACACACCTGGAG (SEQ ID NO: 241)
GCTTGTCTACCTAACCACACCTGGAG (SEQ ID NO: 242)
GCTTGTCTACCTAACTACACCTGGAG (SEQ ID NO: 243)
GCTTGTCTACCTACACACACCTGGAG (SEQ ID NO: 244)
GCTTGTCTACCTACATACACCTGGAG (SEQ ID NO: 245)
GCTTGTCTACCTACCAACACCTGGAG (SEQ ID NO: 246)
GCTTGTCTACCTACCCACACCTGGAG (SEQ ID NO: 247)
GCTTGTCTACCTACCTACACCTGGAG (SEQ ID NO: 248)
GCTTGTCTACCTACGAACACCTGGAG (SEQ ID NO: 249)
GCTTGTCTACCTACGCACACCTGGAG (SEQ ID NO: 250)
GCTTGTCTACCTACTAACACCTGGAG (SEQ ID NO: 251)
GCTTGTCTACCTACTCACACCTGGAG (SEQ ID NO: 252)
GCTTGTCTACCTACTTACACCTGGAG (SEQ ID NO: 253)
GCTTGTCTACCTATCAACACCTGGAG (SEQ ID NO: 254)
GCTTGTCTACCTATCTACACCTGGAG (SEQ ID NO: 255)
GCTTGTCTACCTCAAAACACCTGGAG (SEQ ID NO: 256)
GCTTGTCTACCTCAACACACCTGGAG (SEQ ID NO: 257)
GCTTGTCTACCTCAATACACCTGGAG (SEQ ID NO: 258)
GCTTGTCTACCTCACCACACCTGGAG (SEQ ID NO: 259)
GCTTGTCTACCTCACTACACCTGGAG (SEQ ID NO: 260)
GCTTGTCTACCTCATAACACCTGGAG (SEQ ID NO: 261)
GCTTGTCTACCTCATCACACCTGGAG (SEQ ID NO: 262)
GCTTGTCTACCTCATTACACCTGGAG (SEQ ID NO: 263)
GCTTGTCTACCTCGCAACACCTGGAG (SEQ ID NO: 264)
GCTTGTCTACCTCGCTACACCTGGAG (SEQ ID NO: 265)
GCTTGTCTACCTCTAAACACCTGGAG (SEQ ID NO: 266)
GCTTGTCTACCTCTACACACCTGGAG (SEQ ID NO: 267)
GCTTGTCTACCTCTCAACACCTGGAG (SEQ ID NO: 268)
GCTTGTCTACCTCTCTACACCTGGAG (SEQ ID NO: 269)
GCTTGTCTACCTCTTCACACCTGGAG (SEQ ID NO: 270)
GCTTGTCTACCTGCCAACACCTGGAG (SEQ ID NO: 271)
GCTTGTCTACCTGCCTACACCTGGAG (SEQ ID NO: 272)
GCTTGTCTACCTGCTAACACCTGGAG (SEQ ID NO: 273)
GCTTGTCTACCTGCTTACACCTGGAG (SEQ ID NO: 274)
GCTTGTCTACCTGGAAACACCTGGAG (SEQ ID NO: 275)
GCTTGTCTACCTGTCAACACCTGGAG (SEQ ID NO: 276)
GCTTGTCTACCTGTCTACACCTGGAG (SEQ ID NO: 277)
GCTTGTCTACCTTACCACACCTGGAG (SEQ ID NO: 278)
GCTTGTCTACCTTACTACACCTGGAG (SEQ ID NO: 279)
GCTTGTCTACCTTCAAACACCTGGAG (SEQ ID NO: 280)
GCTTGTCTACCTTCACACACCTGGAG (SEQ ID NO: 281)

FIG. 7 (cont'd)

GCTTGTCTACCTTCATACACCTGGAG (SEQ ID NO: 282)
GCTTGTCTACCTTCTAACACCTGGAG (SEQ ID NO: 283)
GCTTGTCTACCTTCTCACACCTGGAG (SEQ ID NO: 284)
GCTTGTCTACCTTCTTACACCTGGAG (SEQ ID NO: 285)
GCTTGTCTACCTTGTCACACCTGGAG (SEQ ID NO: 286)
GCTTGTCTACGAAAACACACCTGGAG (SEQ ID NO: 287)
GCTTGTCTACGAAACCACACCTGGAG (SEQ ID NO: 288)
GCTTGTCTACGAAACTACACCTGGAG (SEQ ID NO: 289)
GCTTGTCTACGAAATCACACCTGGAG (SEQ ID NO: 290)
GCTTGTCTACGAACACACACCTGGAG (SEQ ID NO: 291)
GCTTGTCTACGAACATACACCTGGAG (SEQ ID NO: 292)
GCTTGTCTACGAACCAACACCTGGAG (SEQ ID NO: 293)
GCTTGTCTACGAACCCACACCTGGAG (SEQ ID NO: 294)
GCTTGTCTACGAACCTACACCTGGAG (SEQ ID NO: 295)
GCTTGTCTACGAACGAACACCTGGAG (SEQ ID NO: 296)
GCTTGTCTACGAACGCACACCTGGAG (SEQ ID NO: 297)
GCTTGTCTACGAACTAACACCTGGAG (SEQ ID NO: 298)
GCTTGTCTACGAACTCACACCTGGAG (SEQ ID NO: 299)
GCTTGTCTACGAACTTACACCTGGAG (SEQ ID NO: 300)
GCTTGTCTACGAATACACACCTGGAG (SEQ ID NO: 301)
GCTTGTCTACGAATCAACACCTGGAG (SEQ ID NO: 302)
GCTTGTCTACGAATCTACACCTGGAG (SEQ ID NO: 303)
GCTTGTCTACGACCAAACACCTGGAG (SEQ ID NO: 304)
GCTTGTCTACGACCACACACCTGGAG (SEQ ID NO: 305)
GCTTGTCTACGACCATACACCTGGAG (SEQ ID NO: 306)
GCTTGTCTACGACCCAACACCTGGAG (SEQ ID NO: 307)
GCTTGTCTACGACCCTACACCTGGAG (SEQ ID NO: 308)
GCTTGTCTACGACCGAACACCTGGAG (SEQ ID NO: 309)
GCTTGTCTACGACCTAACACCTGGAG (SEQ ID NO: 310)
GCTTGTCTACGACCTCACACCTGGAG (SEQ ID NO: 311)
GCTTGTCTACGACCTTACACCTGGAG (SEQ ID NO: 312)
GCTTGTCTACGACGAAACACCTGGAG (SEQ ID NO: 313)
GCTTGTCTACGACGCAACACCTGGAG (SEQ ID NO: 314)
GCTTGTCTACGACGCTACACCTGGAG (SEQ ID NO: 315)
GCTTGTCTACGACGGAACACCTGGAG (SEQ ID NO: 316)
GCTTGTCTACGACTAAACACCTGGAG (SEQ ID NO: 317)
GCTTGTCTACGACTACACACCTGGAG (SEQ ID NO: 318)
GCTTGTCTACGACTCAACACCTGGAG (SEQ ID NO: 319)
GCTTGTCTACGACTCTACACCTGGAG (SEQ ID NO: 320)
GCTTGTCTACGACTTCACACCTGGAG (SEQ ID NO: 321)
GCTTGTCTACGCAAAAACACCTGGAG (SEQ ID NO: 322)
GCTTGTCTACGCAAACACACCTGGAG (SEQ ID NO: 323)
GCTTGTCTACGCAAATACACCTGGAG (SEQ ID NO: 324)
GCTTGTCTACGCAACCACACCTGGAG (SEQ ID NO: 325)
GCTTGTCTACGCAACTACACCTGGAG (SEQ ID NO: 326)
GCTTGTCTACGCAATAACACCTGGAG (SEQ ID NO: 327)
GCTTGTCTACGCAATCACACCTGGAG (SEQ ID NO: 328)

FIG. 7 (cont'd)

GCTTGTCTACGCACACACACCTGGAG (SEQ ID NO: 329)
GCTTGTCTACGCACATACACCTGGAG (SEQ ID NO: 330)
GCTTGTCTACGCACCAACACCTGGAG (SEQ ID NO: 331)
GCTTGTCTACGCACCTACACCTGGAG (SEQ ID NO: 332)
GCTTGTCTACGCACGAACACCTGGAG (SEQ ID NO: 333)
GCTTGTCTACGCACTAACACCTGGAG (SEQ ID NO: 334)
GCTTGTCTACGCACTCACACCTGGAG (SEQ ID NO: 335)
GCTTGTCTACGCACTTACACCTGGAG (SEQ ID NO: 336)
GCTTGTCTACGCATAAACACCTGGAG (SEQ ID NO: 337)
GCTTGTCTACGCATACACACCTGGAG (SEQ ID NO: 338)
GCTTGTCTACGCATCAACACCTGGAG (SEQ ID NO: 339)
GCTTGTCTACGCATCTACACCTGGAG (SEQ ID NO: 340)
GCTTGTCTACGCCAAAACACCTGGAG (SEQ ID NO: 341)
GCTTGTCTACGCCAACACACCTGGAG (SEQ ID NO: 342)
GCTTGTCTACGCCAATACACCTGGAG (SEQ ID NO: 343)
GCTTGTCTACGCCACTACACCTGGAG (SEQ ID NO: 344)
GCTTGTCTACGCCATAACACCTGGAG (SEQ ID NO: 345)
GCTTGTCTACGCCATCACACCTGGAG (SEQ ID NO: 346)
GCTTGTCTACGCCATTACACCTGGAG (SEQ ID NO: 347)
GCTTGTCTACGCCCAAACACCTGGAG (SEQ ID NO: 348)
GCTTGTCTACGCCCATACACCTGGAG (SEQ ID NO: 349)
GCTTGTCTACGCCCTAACACCTGGAG (SEQ ID NO: 350)
GCTTGTCTACGCCCTTACACCTGGAG (SEQ ID NO: 351)
GCTTGTCTACGCCTAAACACCTGGAG (SEQ ID NO: 352)
GCTTGTCTACGCCTACACACCTGGAG (SEQ ID NO: 353)
GCTTGTCTACGCCTCAACACCTGGAG (SEQ ID NO: 354)
GCTTGTCTACGCCTCTACACCTGGAG (SEQ ID NO: 355)
GCTTGTCTACGCCTTCACACCTGGAG (SEQ ID NO: 356)
GCTTGTCTACGCTAAAACACCTGGAG (SEQ ID NO: 357)
GCTTGTCTACGCTAACACACCTGGAG (SEQ ID NO: 358)
GCTTGTCTACGCTACCACACCTGGAG (SEQ ID NO: 359)
GCTTGTCTACGCTACTACACCTGGAG (SEQ ID NO: 360)
GCTTGTCTACGCTATCACACCTGGAG (SEQ ID NO: 361)
GCTTGTCTACGCTCAAACACCTGGAG (SEQ ID NO: 362)
GCTTGTCTACGCTCACACACCTGGAG (SEQ ID NO: 363)
GCTTGTCTACGCTCATACACCTGGAG (SEQ ID NO: 364)
GCTTGTCTACGCTCTAACACCTGGAG (SEQ ID NO: 365)
GCTTGTCTACGCTCTCACACCTGGAG (SEQ ID NO: 366)
GCTTGTCTACGCTCTTACACCTGGAG (SEQ ID NO: 367)
GCTTGTCTACGCTGCTACACCTGGAG (SEQ ID NO: 368)
GCTTGTCTACGCTGGAACACCTGGAG (SEQ ID NO: 369)
GCTTGTCTACGCTGTCACACCTGGAG (SEQ ID NO: 370)
GCTTGTCTACGCTTACACACCTGGAG (SEQ ID NO: 371)
GCTTGTCTACGCTTCAACACCTGGAG (SEQ ID NO: 372)
GCTTGTCTACGCTTCTACACCTGGAG (SEQ ID NO: 373)
GCTTGTCTACGGAAAAACACCTGGAG (SEQ ID NO: 374)
GCTTGTCTACGGAAACACACCTGGAG (SEQ ID NO: 375)

FIG. 7 (cont'd)

GCTTGTCTACGGAAATACACCTGGAG (SEQ ID NO: 376)
GCTTGTCTACGGAACCACACCTGGAG (SEQ ID NO: 377)
GCTTGTCTACGGAACTACACCTGGAG (SEQ ID NO: 378)
GCTTGTCTACGGAATAACACCTGGAG (SEQ ID NO: 379)
GCTTGTCTACGGAATCACACCTGGAG (SEQ ID NO: 380)
GCTTGTCTACGGACCAACACCTGGAG (SEQ ID NO: 381)
GCTTGTCTACGGACCTACACCTGGAG (SEQ ID NO: 382)
GCTTGTCTACGGACGAACACCTGGAG (SEQ ID NO: 383)
GCTTGTCTACGGACTAACACCTGGAG (SEQ ID NO: 384)
GCTTGTCTACGGACTCACACCTGGAG (SEQ ID NO: 385)
GCTTGTCTACGGACTTACACCTGGAG (SEQ ID NO: 386)
GCTTGTCTACGGCTAAACACCTGGAG (SEQ ID NO: 387)
GCTTGTCTACGGCTACACACCTGGAG (SEQ ID NO: 388)
GCTTGTCTACGGCTCAACACCTGGAG (SEQ ID NO: 389)
GCTTGTCTACGGCTTCACACCTGGAG (SEQ ID NO: 390)
GCTTGTCTACTAAACCACACCTGGAG (SEQ ID NO: 391)
GCTTGTCTACTAACACACACCTGGAG (SEQ ID NO: 392)
GCTTGTCTACTAACCAACACCTGGAG (SEQ ID NO: 393)
GCTTGTCTACTAACCCACACCTGGAG (SEQ ID NO: 394)
GCTTGTCTACTAACCTACACCTGGAG (SEQ ID NO: 395)
GCTTGTCTACTAACGAACACCTGGAG (SEQ ID NO: 396)
GCTTGTCTACTAACGCACACCTGGAG (SEQ ID NO: 397)
GCTTGTCTACTAACTCACACCTGGAG (SEQ ID NO: 398)
GCTTGTCTACTACACCACACCTGGAG (SEQ ID NO: 399)
GCTTGTCTACTACACTACACCTGGAG (SEQ ID NO: 400)
GCTTGTCTACTACATCACACCTGGAG (SEQ ID NO: 401)
GCTTGTCTACTACCAAACACCTGGAG (SEQ ID NO: 402)
GCTTGTCTACTACCACACACCTGGAG (SEQ ID NO: 403)
GCTTGTCTACTACCATACACCTGGAG (SEQ ID NO: 404)
GCTTGTCTACTACCCAACACCTGGAG (SEQ ID NO: 405)
GCTTGTCTACTACCCCACACCTGGAG (SEQ ID NO: 406)
GCTTGTCTACTACCCTACACCTGGAG (SEQ ID NO: 407)
GCTTGTCTACTACCGAACACCTGGAG (SEQ ID NO: 408)
GCTTGTCTACTACCGCACACCTGGAG (SEQ ID NO: 409)
GCTTGTCTACTACCTAACACCTGGAG (SEQ ID NO: 410)
GCTTGTCTACTACCTCACACCTGGAG (SEQ ID NO: 411)
GCTTGTCTACTACCTTACACCTGGAG (SEQ ID NO: 412)
GCTTGTCTACTACGAAACACCTGGAG (SEQ ID NO: 413)
GCTTGTCTACTACGCAACACCTGGAG (SEQ ID NO: 414)
GCTTGTCTACTACGCCACACCTGGAG (SEQ ID NO: 415)
GCTTGTCTACTACGCTACACCTGGAG (SEQ ID NO: 416)
GCTTGTCTACTACGGAACACCTGGAG (SEQ ID NO: 417)
GCTTGTCTACTACTACACACCTGGAG (SEQ ID NO: 418)
GCTTGTCTACTACTCAACACCTGGAG (SEQ ID NO: 419)
GCTTGTCTACTACTCTACACCTGGAG (SEQ ID NO: 420)
GCTTGTCTACTACTTCACACCTGGAG (SEQ ID NO: 421)
GCTTGTCTACTATCACACACCTGGAG (SEQ ID NO: 422)

FIG. 7 (cont'd)

GCTTGTCTACTATCCCACACCTGGAG (SEQ ID NO: 423)
GCTTGTCTACTATCCTACACCTGGAG (SEQ ID NO: 424)
GCTTGTCTACTATCGCACACCTGGAG (SEQ ID NO: 425)
GCTTGTCTACTATCTCACACCTGGAG (SEQ ID NO: 426)
GCTTGTCTACTCAAACACACCTGGAG (SEQ ID NO: 427)
GCTTGTCTACTCAACCACACCTGGAG (SEQ ID NO: 428)
GCTTGTCTACTCAACTACACCTGGAG (SEQ ID NO: 429)
GCTTGTCTACTCAATCACACCTGGAG (SEQ ID NO: 430)
GCTTGTCTACTCACACACACCTGGAG (SEQ ID NO: 431)
GCTTGTCTACTCACATACACCTGGAG (SEQ ID NO: 432)
GCTTGTCTACTCACCAACACCTGGAG (SEQ ID NO: 433)
GCTTGTCTACTCACCCACACCTGGAG (SEQ ID NO: 434)
GCTTGTCTACTCACCTACACCTGGAG (SEQ ID NO: 435)
GCTTGTCTACTCACGAACACCTGGAG (SEQ ID NO: 436)
GCTTGTCTACTCACGCACACCTGGAG (SEQ ID NO: 437)
GCTTGTCTACTCACTAACACCTGGAG (SEQ ID NO: 438)
GCTTGTCTACTCACTCACACCTGGAG (SEQ ID NO: 439)
GCTTGTCTACTCACTTACACCTGGAG (SEQ ID NO: 440)
GCTTGTCTACTCATACACACCTGGAG (SEQ ID NO: 441)
GCTTGTCTACTCATCAACACCTGGAG (SEQ ID NO: 442)
GCTTGTCTACTCATCTACACCTGGAG (SEQ ID NO: 443)
GCTTGTCTACTCGCAAACACCTGGAG (SEQ ID NO: 444)
GCTTGTCTACTCGCACACACCTGGAG (SEQ ID NO: 445)
GCTTGTCTACTCGCATACACCTGGAG (SEQ ID NO: 446)
GCTTGTCTACTCGCCAACACCTGGAG (SEQ ID NO: 447)
GCTTGTCTACTCGCCTACACCTGGAG (SEQ ID NO: 448)
GCTTGTCTACTCGCTAACACCTGGAG (SEQ ID NO: 449)
GCTTGTCTACTCGCTCACACCTGGAG (SEQ ID NO: 450)
GCTTGTCTACTCGCTTACACCTGGAG (SEQ ID NO: 451)
GCTTGTCTACTCTAACACACCTGGAG (SEQ ID NO: 452)
GCTTGTCTACTCTACCACACCTGGAG (SEQ ID NO: 453)
GCTTGTCTACTCTACTACACCTGGAG (SEQ ID NO: 454)
GCTTGTCTACTCTATCACACCTGGAG (SEQ ID NO: 455)
GCTTGTCTACTCTCAAACACCTGGAG (SEQ ID NO: 456)
GCTTGTCTACTCTCACACACCTGGAG (SEQ ID NO: 457)
GCTTGTCTACTCTCATACACCTGGAG (SEQ ID NO: 458)
GCTTGTCTACTCTCGCACACCTGGAG (SEQ ID NO: 459)
GCTTGTCTACTCTCTAACACCTGGAG (SEQ ID NO: 460)
GCTTGTCTACTCTCTCACACCTGGAG (SEQ ID NO: 461)
GCTTGTCTACTCTCTTACACCTGGAG (SEQ ID NO: 462)
GCTTGTCTACTCTGCCACACCTGGAG (SEQ ID NO: 463)
GCTTGTCTACTCTGCTACACCTGGAG (SEQ ID NO: 464)
GCTTGTCTACTCTGGAACACCTGGAG (SEQ ID NO: 465)
GCTTGTCTACTCTGTCACACCTGGAG (SEQ ID NO: 466)
GCTTGTCTACTCTTACACACCTGGAG (SEQ ID NO: 467)
GCTTGTCTACTCTTCAACACCTGGAG (SEQ ID NO: 468)
GCTTGTCTACTCTTCTACACCTGGAG (SEQ ID NO: 469)

FIG. 7 (cont'd)

GCTTGTCTACTGCCAAACACCTGGAG (SEQ ID NO: 470)
GCTTGTCTACTGCCACACACCTGGAG (SEQ ID NO: 471)
GCTTGTCTACTGCCATACACCTGGAG (SEQ ID NO: 472)
GCTTGTCTACTGCCCAACACCTGGAG (SEQ ID NO: 473)
GCTTGTCTACTGCCCTACACCTGGAG (SEQ ID NO: 474)
GCTTGTCTACTGCCTAACACCTGGAG (SEQ ID NO: 475)
GCTTGTCTACTGCCTCACACCTGGAG (SEQ ID NO: 476)
GCTTGTCTACTGCCTTACACCTGGAG (SEQ ID NO: 477)
GCTTGTCTACTGCTAAACACCTGGAG (SEQ ID NO: 478)
GCTTGTCTACTGCTACACACCTGGAG (SEQ ID NO: 479)
GCTTGTCTACTGCTCAACACCTGGAG (SEQ ID NO: 480)
GCTTGTCTACTGCTCTACACCTGGAG (SEQ ID NO: 481)
GCTTGTCTACTGCTTCACACCTGGAG (SEQ ID NO: 482)
GCTTGTCTACTGGAAAACACCTGGAG (SEQ ID NO: 483)
GCTTGTCTACTGGAACACACCTGGAG (SEQ ID NO: 484)
GCTTGTCTACTGGAATACACCTGGAG (SEQ ID NO: 485)
GCTTGTCTACTGGACCACACCTGGAG (SEQ ID NO: 486)
GCTTGTCTACTGGACTACACCTGGAG (SEQ ID NO: 487)
GCTTGTCTACTGTCAAACACCTGGAG (SEQ ID NO: 488)
GCTTGTCTACTGTCACACACCTGGAG (SEQ ID NO: 489)
GCTTGTCTACTGTCATACACCTGGAG (SEQ ID NO: 490)
GCTTGTCTACTGTCTAACACCTGGAG (SEQ ID NO: 491)
GCTTGTCTACTGTCTCACACCTGGAG (SEQ ID NO: 492)
GCTTGTCTACTGTCTTACACCTGGAG (SEQ ID NO: 493)
GCTTGTCTACTTACACACACCTGGAG (SEQ ID NO: 494)
GCTTGTCTACTTACCAACACCTGGAG (SEQ ID NO: 495)
GCTTGTCTACTTACCCACACCTGGAG (SEQ ID NO: 496)
GCTTGTCTACTTACCTACACCTGGAG (SEQ ID NO: 497)
GCTTGTCTACTTACGAACACCTGGAG (SEQ ID NO: 498)
GCTTGTCTACTTACGCACACCTGGAG (SEQ ID NO: 499)
GCTTGTCTACTTACTCACACCTGGAG (SEQ ID NO: 500)
GCTTGTCTACTTCAACACACCTGGAG (SEQ ID NO: 501)
GCTTGTCTACTTCACCACACCTGGAG (SEQ ID NO: 502)
GCTTGTCTACTTCACTACACCTGGAG (SEQ ID NO: 503)
GCTTGTCTACTTCATCACACCTGGAG (SEQ ID NO: 504)
GCTTGTCTACTTCTACACACCTGGAG (SEQ ID NO: 505)
GCTTGTCTACTTCTCAACACCTGGAG (SEQ ID NO: 506)
GCTTGTCTACTTCTCTACACCTGGAG (SEQ ID NO: 507)
GCTTGTCTACTTCTTCACACCTGGAG (SEQ ID NO: 508)
GCTTGTCTACTTGTCAACACCTGGAG (SEQ ID NO: 509)
GCTTGTCTACTTGTCTACACCTGGAG (SEQ ID NO: 510)

FIG. 8

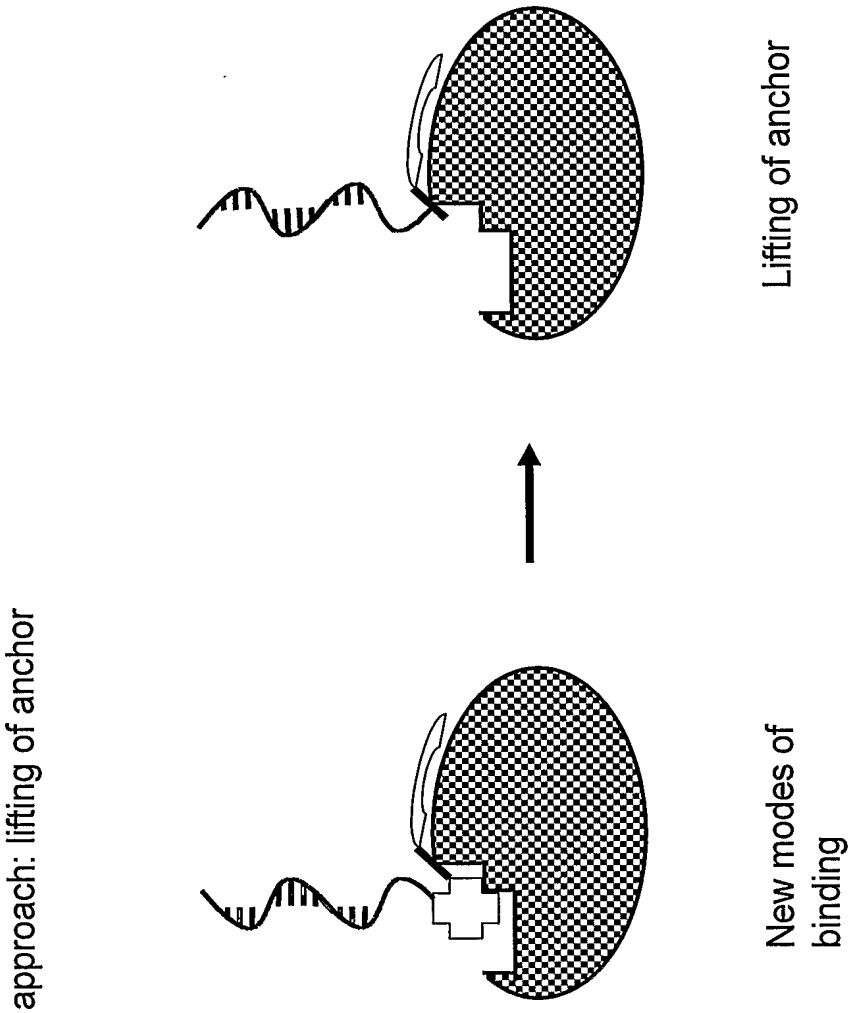

Anchor-based approach: lifting of anchor
New modes of binding
Lifting of anchor

FIG. 9

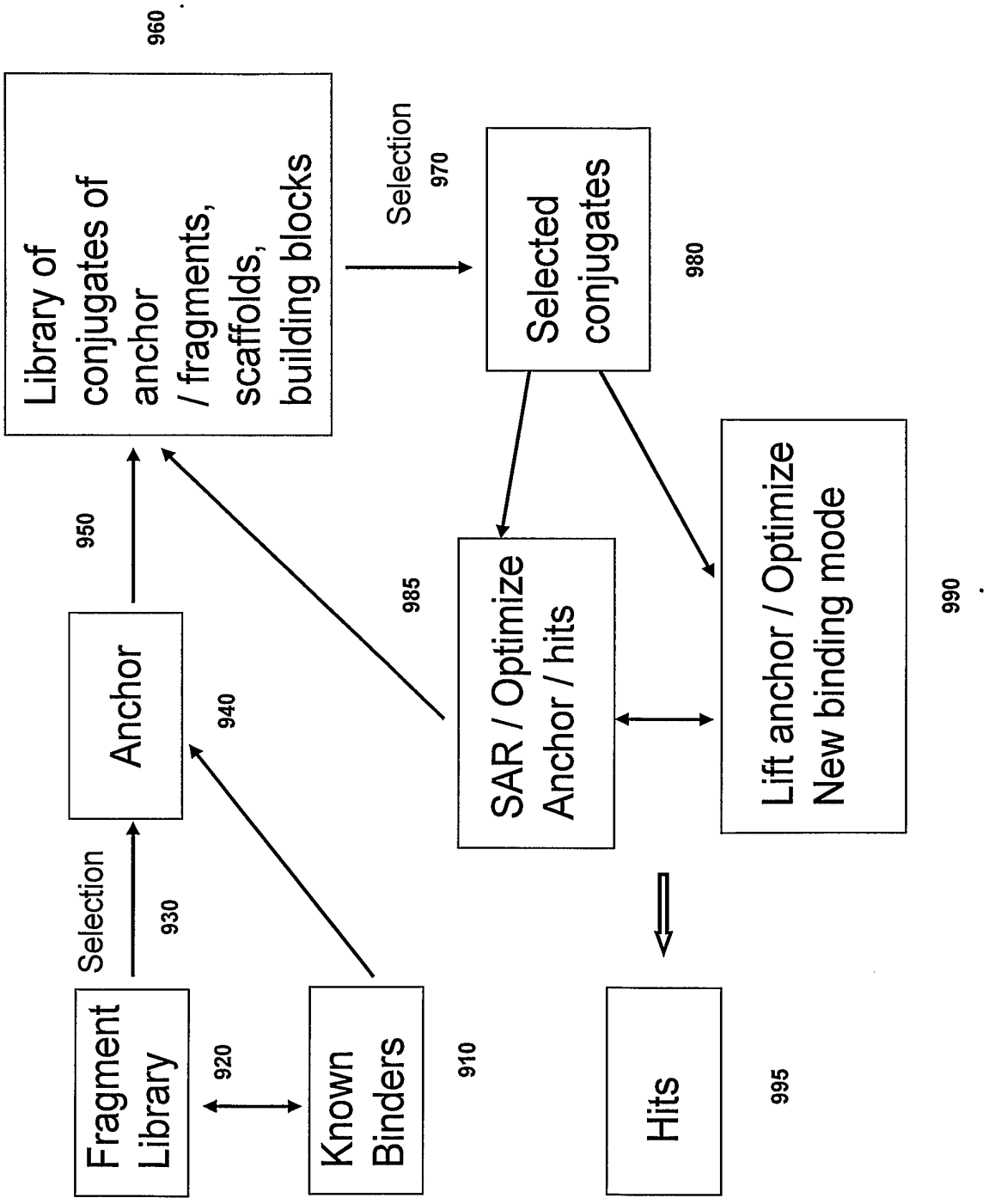

Library of conjugates of anchor / fragments, scaffolds, building blocks
960
Selection
970
Selected conjugates
980
950
Anchor
940
985
SAR / Optimize Anchor / hits
Lift anchor / Optimize New binding mode
990
Selection
930
Fragment Library
920
Known Binders
910
Hits
995

FIG. 10

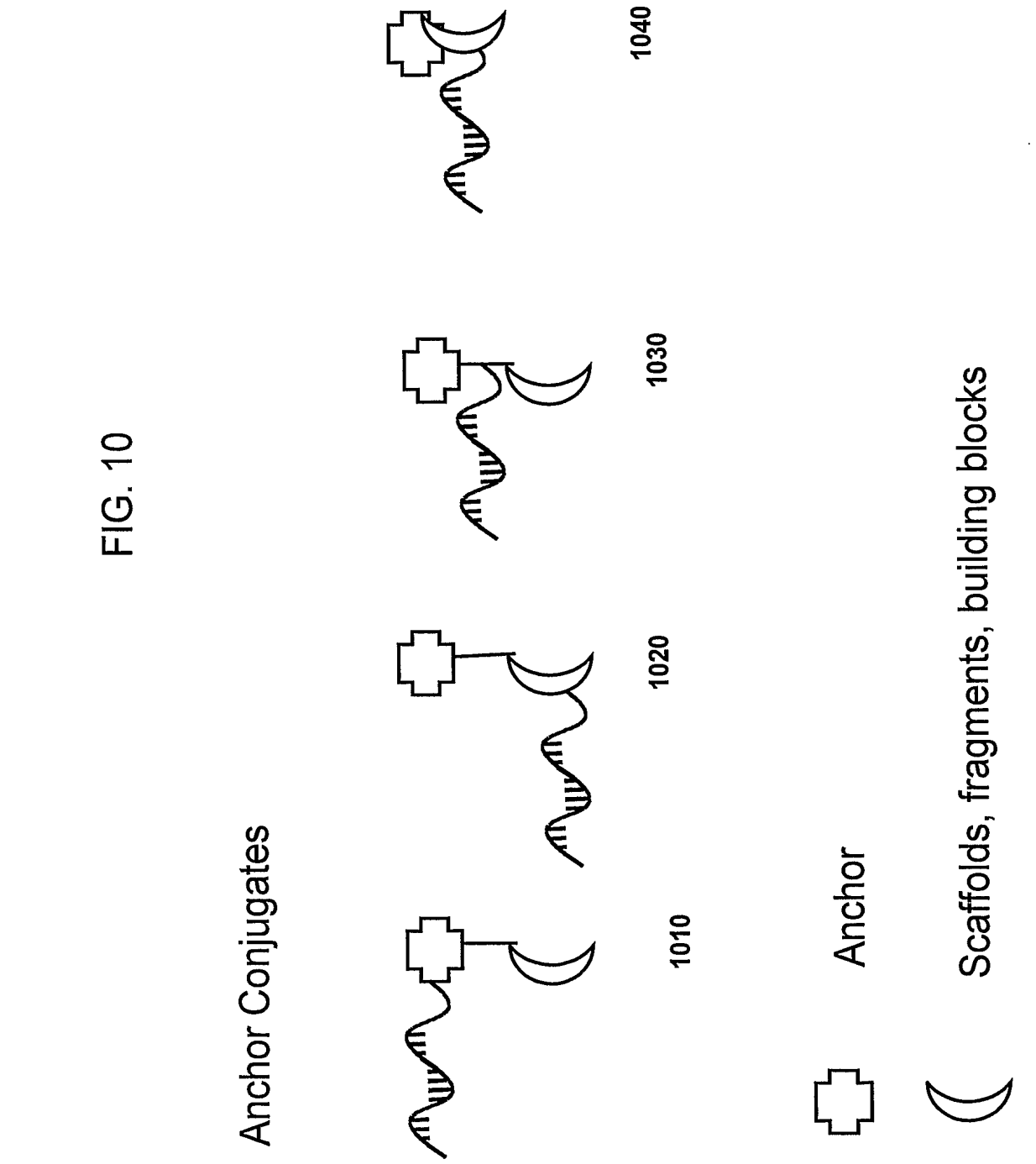
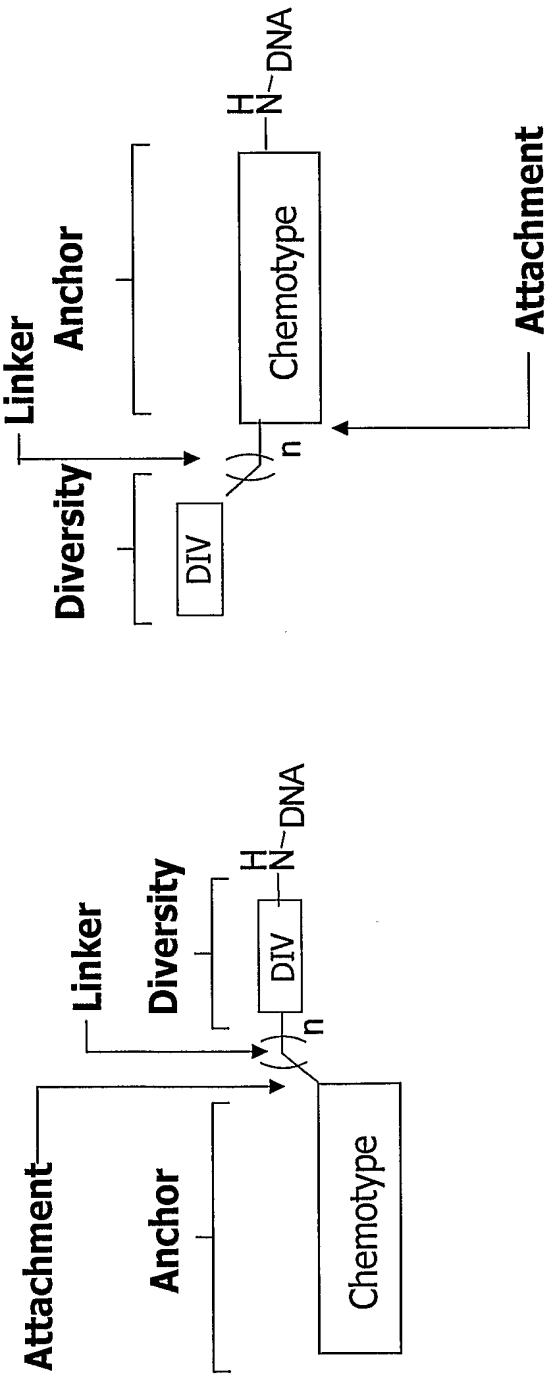

Anchor Conjugates
1010
1020
1030
1040
Anchor
Scaffolds, fragments, building blocks

FIG. 11



Number of Compounds = 1 Anchor x N_x Attachments x N_y Linkers x N_z Diversity
N_x = 1-10; N_y = 1-10; N_z = 10-20,000
Total = 45,000 when N_x = 3; N_y = 3; N_z = 5,000

FIG. 12

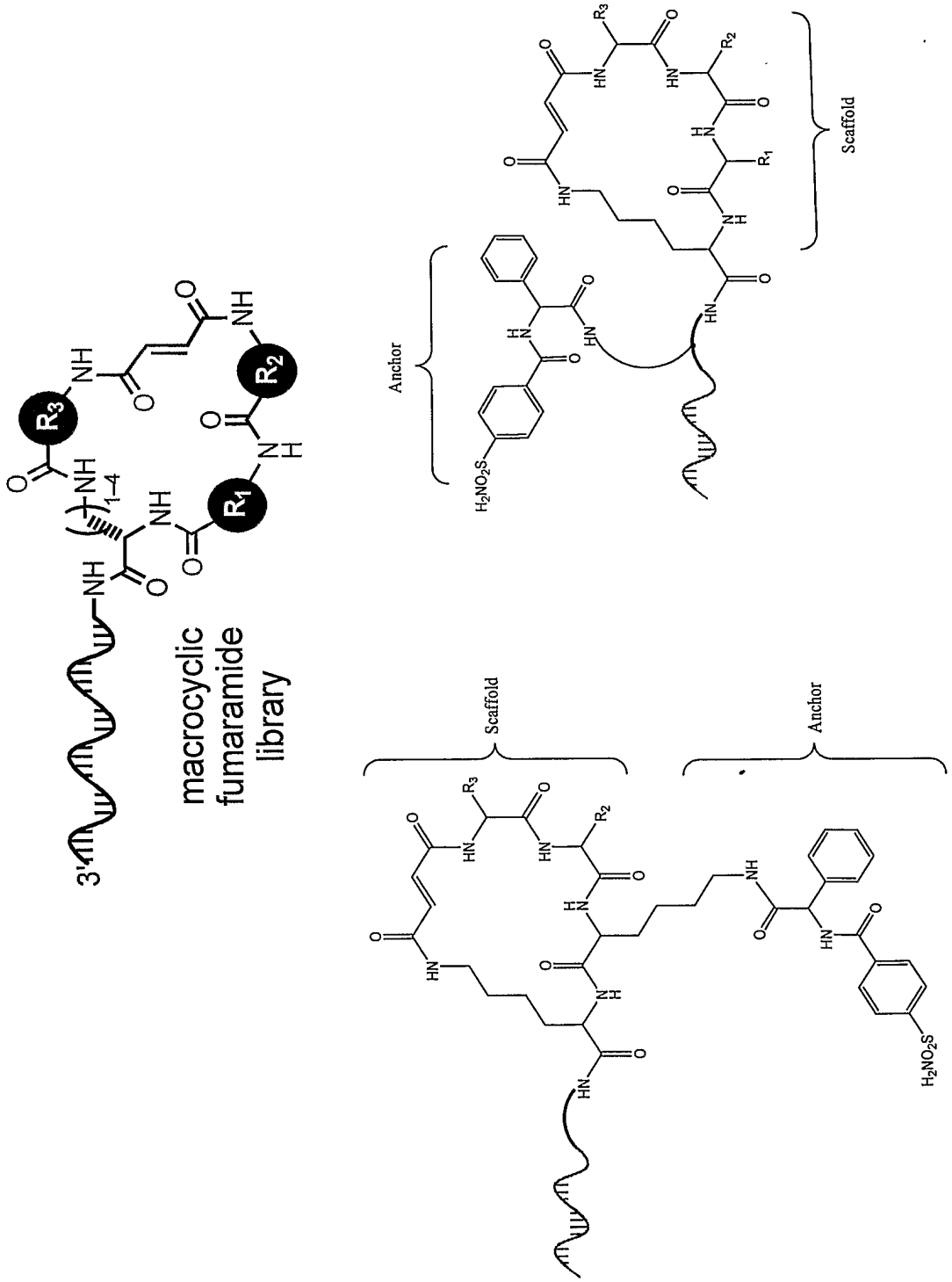

macrocyclic
fumaramide
library
Anchor
Scaffold
Scaffold
Anchor

FIG. 13

Architecture of 3'-Amino-Modifier C7 DNA Conjugate

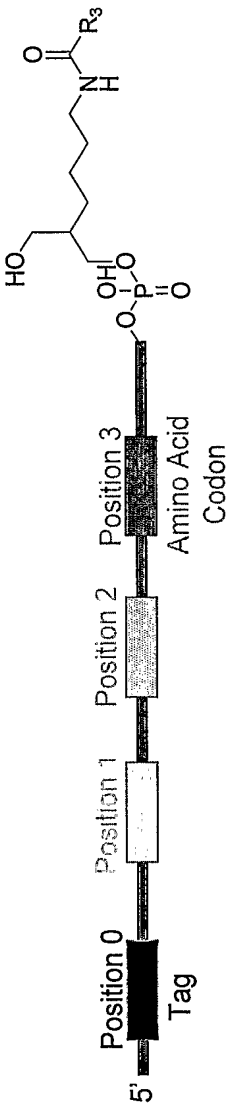


FIG. 14

See FIG. 15 for exemplary R_3's

Architecture of 5'-Amino-Modifier 5 DNA Conjugate

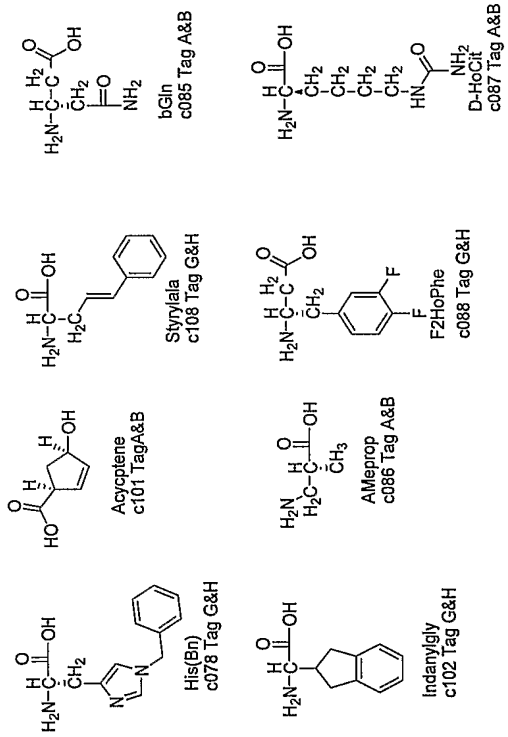

His(Bn)
c078 Tag G&H
Acyclptene
c101 TagA&B
Styrylala
c108 Tag G&H
bGln
c085 Tag A&B
Indanylgly
c102 Tag G&H
AMeprop
c086 Tag A&B
F2HoPhe
c088 Tag G&H
D-HoCit
c087 Tag A&B
Freidam
c065 Tag G&H
LysFor
c108 Tag A&B
D or L Bip
c0498&c112 Tag G&H
Leu
c086 Tag G&H
Valeram
c049 Tag A&B
ABztB
c101 Tag G&H
ArgMe2
c065 Tag A&B
Cxhca
c088 Tag A&B

FIG. 15

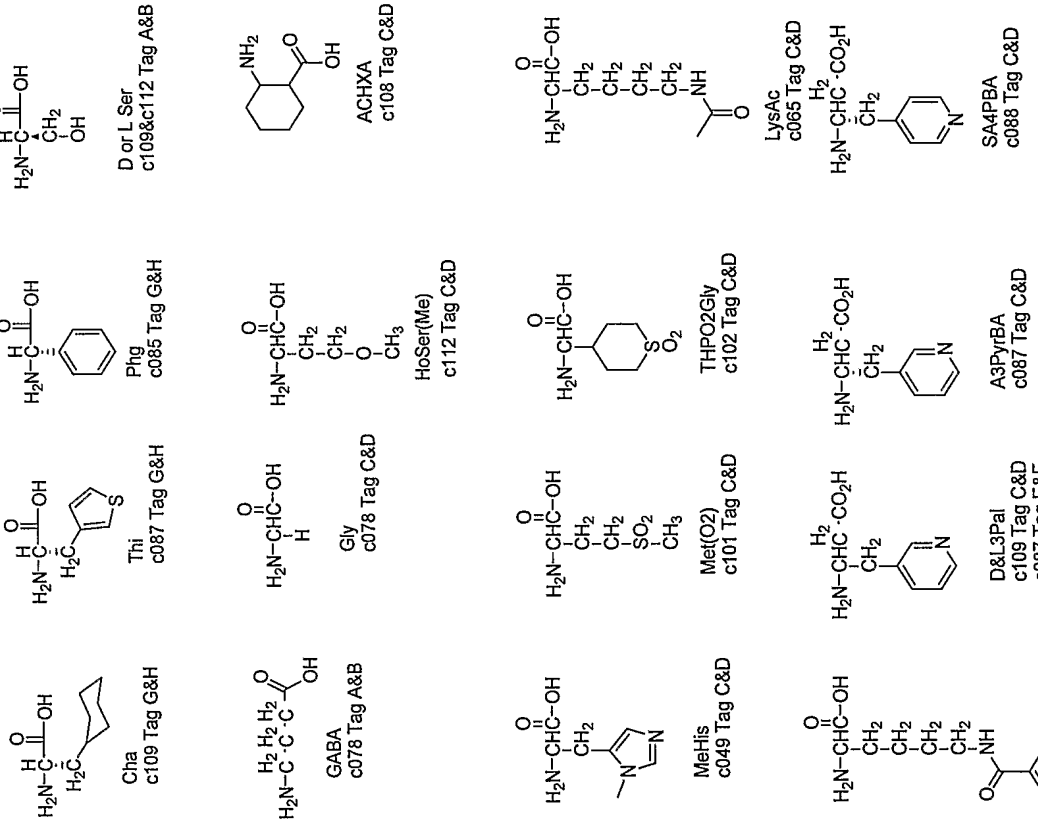

Cha
c109 Tag G&H
Thi
c087 Tag G&H
Phg
c085 Tag G&H
D or L Ser
c109&c112 Tag A&B
GABA
c078 Tag A&B
Gly
c078 Tag C&D
HoSer(Me)
c112 Tag C&D
ACHXA
c108 Tag C&D
MeHis
c049 Tag C&D
Met(O2)
c101 Tag C&D
THPO2Gly
c102 Tag C&D
LysAc
c065 Tag C&D
Lys(Nic)
c085 Tag C&D
D&L3Pal
c109 Tag C&D
c087 Tag E&F
A3PyrBA
c087 Tag C&D
SA4PBA
c088 Tag C&D

FIG. 15
(cont'd)

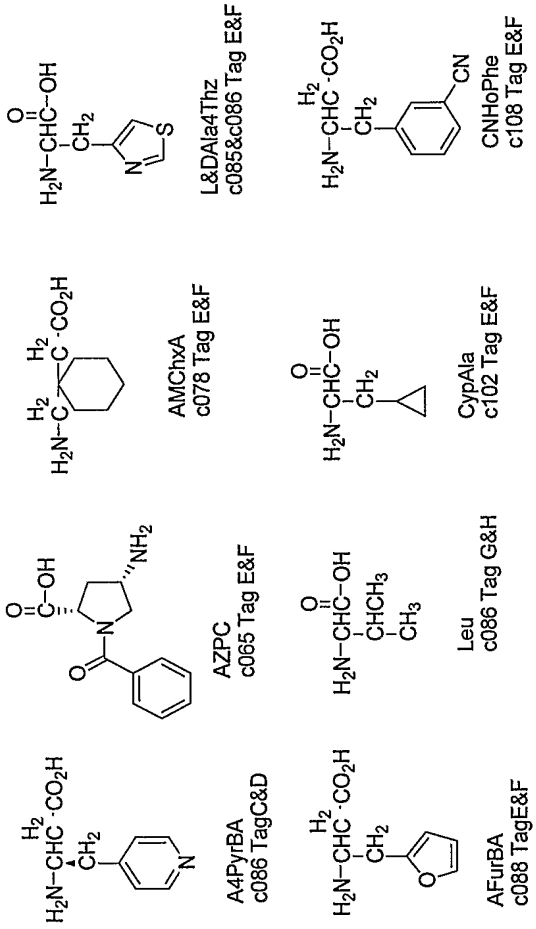
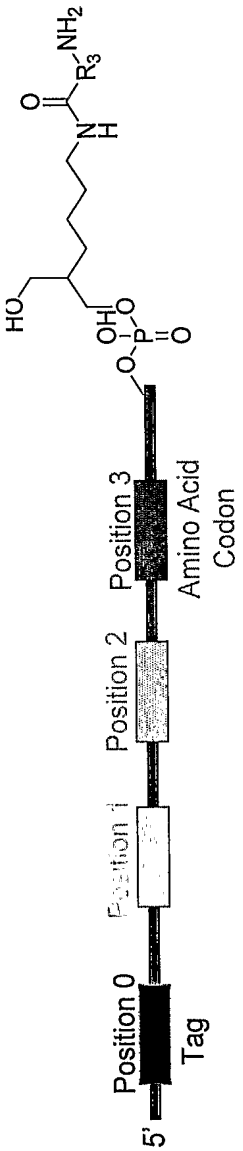

A4PyrBA
c086 TagC&D
AZPC
c065 Tag E&F
AMChxA
c078 Tag E&F
L&DAla4Thz
c085&c086 Tag E&F
AFurBA
c088 TagE&F
Leu
c086 Tag G&H
CypAla
c102 Tag E&F
CNHoPhe
c108 Tag E&F
AThiBA
DiMeOPhe
Ala2Fur

FIG. 15
(cont'd)

FIG. 16



Architecture of an exemplary DPC fragments
R_3 is the structural diversity represented by the sample Target Binding Elements (FIG. 15)

FIG. 17

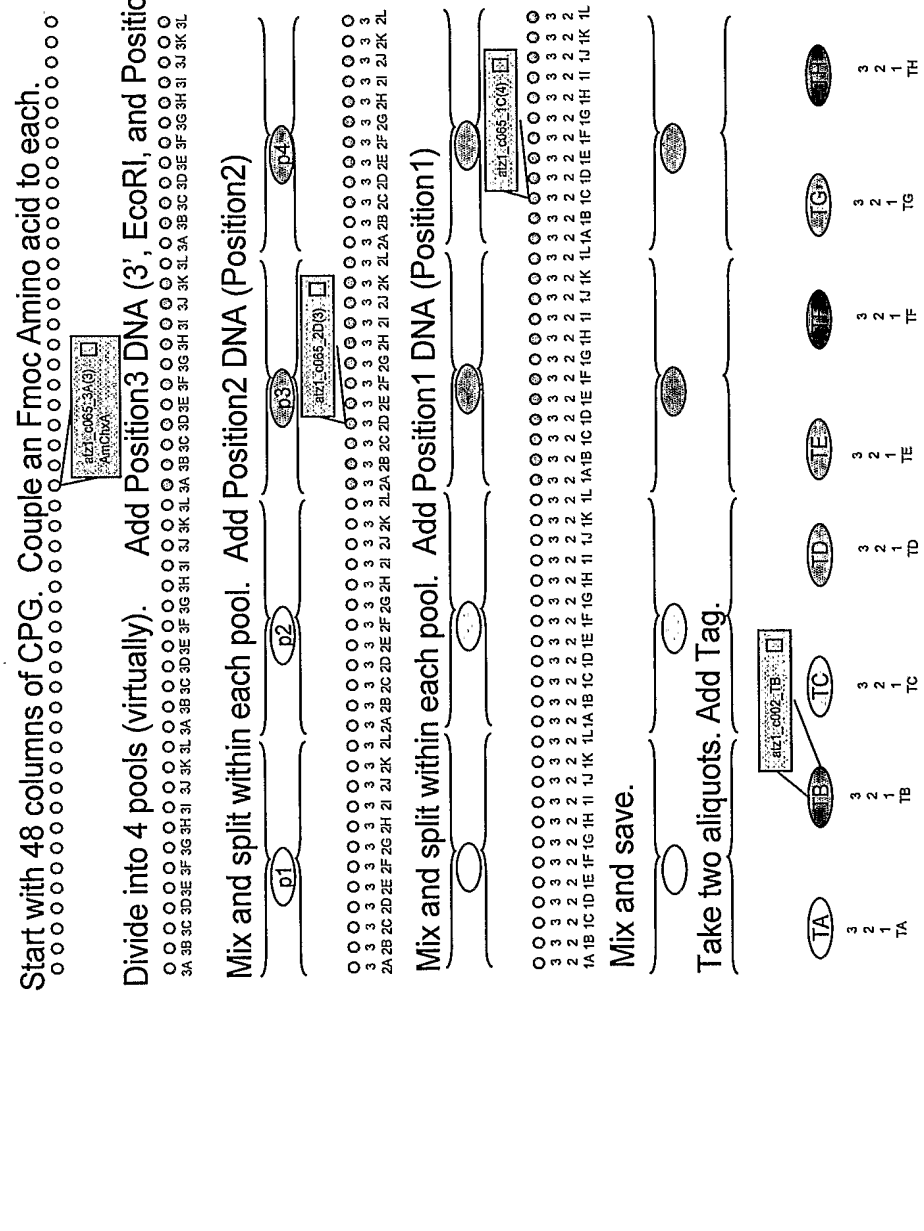
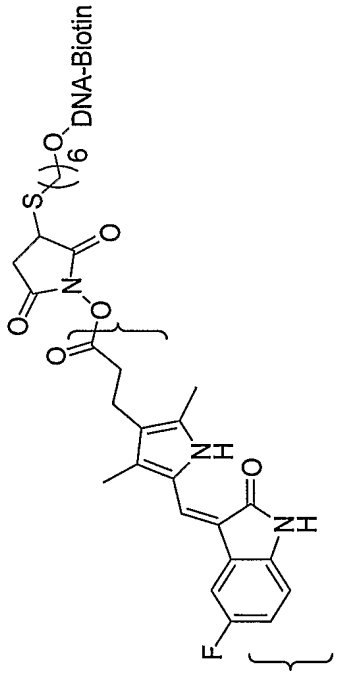

Start with 48 columns of CPG. Couple an Fmoc Amino acid to each.
Divide into 4 pools (virtually). Add Position3 DNA (3', EcoRI, and Position3)
Mix and split within each pool. Add Position2 DNA (Position2)
Mix and split within each pool. Add Position1 DNA (Position1)
Mix and save.
Take two aliquots. Add Tag.
8x3 col. vol.
Mix and split strategy for DPC fragments

FIG. 18



FOPP

Structure of an example of a FOPP-labeled Fragment Conjugate
FOPP: Name from (Z)-3-(5-((5-fluoro-2-oxoindolin-3-ylidene)methyl)-2,4-dimethyl-1*H*-pyrrol-3-yl)propanoic acid

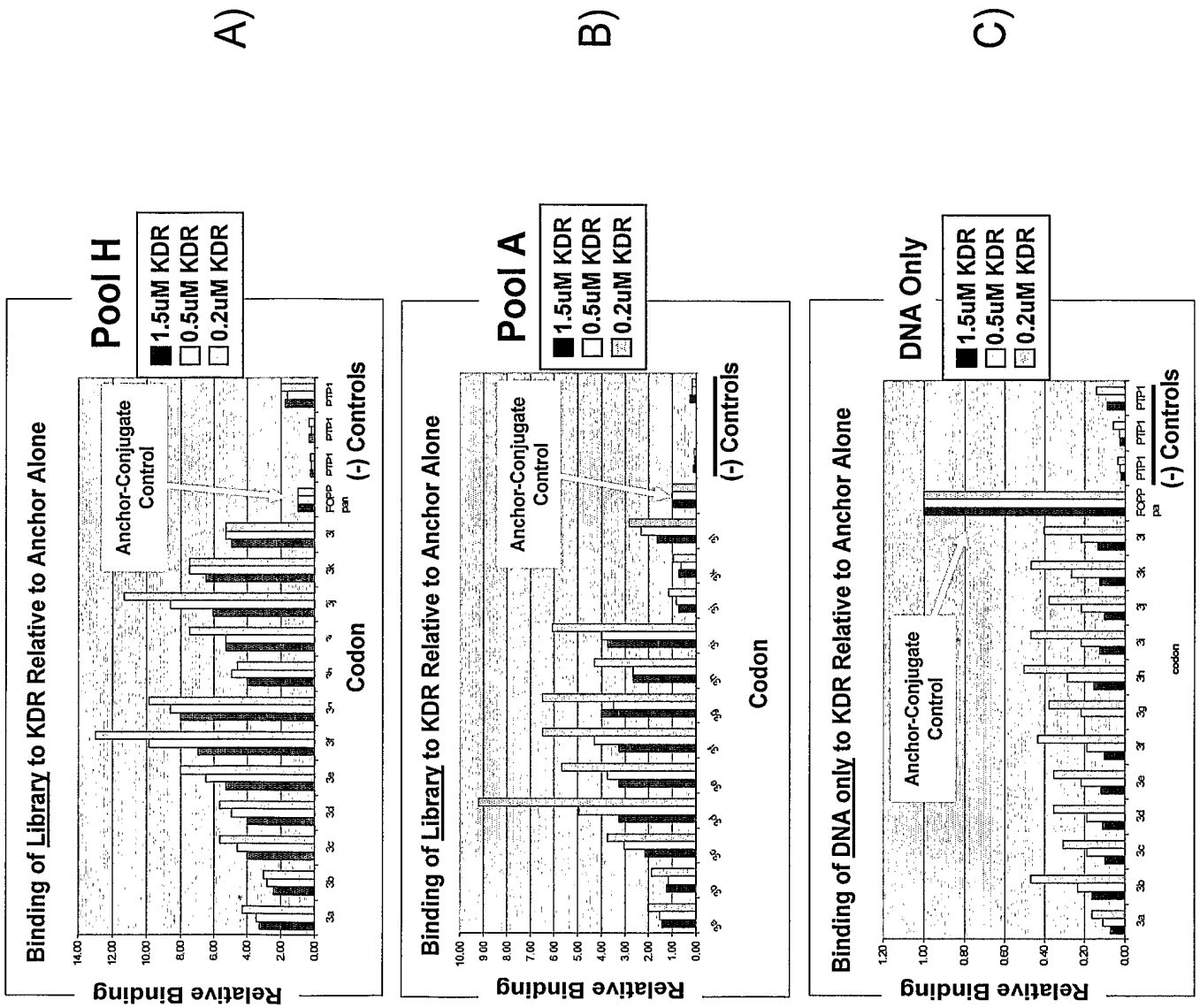

A)
Pool H
1.5uM KDR
0.5uM KDR
0.2uM KDR
Binding of Library to KDR Relative to Anchor Alone
Anchor-Conjugate Control
Relative Binding
Codon
(-) Controls
B)
Pool A
1.5uM KDR
0.5uM KDR
0.2uM KDR
Binding of Library to KDR Relative to Anchor Alone
Anchor-Conjugate Control
Relative Binding
Codon
(-) Controls
C)
DNA Only
1.5uM KDR
0.5uM KDR
0.2uM KDR
Binding of DNA only to KDR Relative to Anchor Alone
Anchor-Conjugate Control
Relative Binding
codon
(-) Controls

FIG. 19